Open Science in Conservation: Combining Citizen Science and Remote Sensing Approaches for Habitat Monitoring

by

Samantha Suter

BSc Zoology

MSc Biodiversity and Conservation

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

October 2023

School of Social and Environmental Sustainability

College of Social Sciences

University of Glasgow

"This is renaissance, your dentist now an authority on butterflies and you (in retrospect this happened so pleasantly, watching clouds one afternoon) connected by Twitter to the National Weather Service. This is revolution, breaking down barriers between expert and amateur, with new collaborations across class and education. Pygmy hunters and gatherers use smartphones to document deforestation in the Congo Basin. High school students identify fossils in soils from ancient seas in upstate New York. Do-it-yourself biologists make centrifuges at home. This is falling in love with the world, and this is science, and at the risk of sounding too much an idealist, I have come to believe they are the same thing".

(Russell 2014, p. 11)

Abstract

Global biodiversity conservation efforts have not been sufficient to reverse declining biodiversity trends. These shortcomings have led to gaps in monitoring initiatives for habitats, taxa, and regions. Public disconnect from nature and a lack of policy implementation to action scientific research have exacerbated these issues. Alternative monitoring tools, such as citizen science (CS) and remote sensing (RS), have the potential to increase the spatial and temporal reach of monitoring, meeting numerous global biodiversity targets. As such, there are calls for CS and RS to be united. Therefore, this thesis aimed to pair CS and RS and a provide a low cost, low intensity open science (OS) tool in Scotland. This tool sought to map a UK priority habitat – species-rich grasslands (SRGs), which have been widely reduced in area, resulting in coupled invertebrate decline and diminished ecosystem functioning. To increase success of monitoring attempts, an OS approach was adopted, whereby collaboration with stakeholders was key to enhance research impact and scientific democratisation.

To establish an OS framework, openness in current biodiversity monitoring CS surveys was initially investigated, revealing that these surveys did not consistently adhere to OS practices. This research was vital to informing the design of the CS survey of this thesis and identifying efforts to ensure full openness along the scientific process. In determining OS practices, open access Sentinel-2 satellite imagery was acquired to create a habitat classification model, with a final accuracy of 98.6%. Other RS applications were explored, such as the Spectral Variation Hypothesis and grassland trait retrieval, to investigate open access RS data in subsequent mapping attempts of SRGs. The results found no significant relationship between spectral and species diversity, and grassland traits were mostly poorly predicted across spatial and spectral scales. The habitat prediction model was applied to satellite imagery across Scotland, predicting areas of SRGs. In exploring the model outputs, a CS survey, Ecosystem Explorers, was created with Butterfly Conservation, where participants surveyed these predicted areas. Data on previously surveyed SRGs was provided by collaborators such as NatureScot, Plantlife, and the Botanical Society of Britain and Ireland. The model predicted and citizen ground-truthed SRG locations were compared and a poor alignment of 17.65% was found. Participant identification experience and habitat assessment confidence did not affect the level of agreement between the model predictions and ground-truthed observations. However, OS was implemented to combine CS and RS in a highly accessible project, with a predicted open score of 0.92/1.

The thesis provides an example of a novel, OS biodiversity monitoring tool by combining CS and RS methods. The attempt to predict SRGs across Scotland utilising this instrument was unsuccessful. Although the cause of this poor alignment is unclear, citizen scientists appear to be equally as consistent as professionals in their observations, suggesting potential weaknesses in the national application of the model for SRG predictions. This would need further exploration and then poses the question of how an OS tool can be used to improve biodiversity monitoring. Throughout the thesis, guidance and recommendations are provided on how and where OS practices and methodologies can be used for future biodiversity monitoring.

Table of Contents

List of Tables

List of Figures

Acknowledgements

Expressing gratitude at this level is near impossible to put into words but I will try, nonetheless. I will forever be indebted to the many people who have supported and encouraged me throughout my PhD journey, without which I may not be writing this now.

To my parents, David and Petra, who have been my number one cheerleaders, boasting about me and pretending to understand my research; I could not have finished without their continuous care and refuge when circumstances became too much. To my brother, Kyle, whose constant contesting is always great motivation. And of course, to my cats, Alfie and Peppa, my forever happy pills.

To my supervisors, Natalie Welden and Brian Barrett, who probably cannot wait to never receive essays parading as emails from me again. I could not be more thankful to have had supervisors who have not only been incredibly involved, but for the friendly environment that they created. For their expertise, advice, and the experiences I have had working with them, I am overwhelmingly appreciative. I would also like to thank Kenny Roberts for accompanying me around Scotland, trusting me with extortionately expensive equipment, and his knowledge and dedication to helping me. I owe a lot to Michael Muir, who I thank immensely for his help in the lab, emotional support chocolate provision, and the laughs along the way. A big thank you to Steven Gillespie also for the opportunities he extended to me. I hope I have the chance to work with all of them again.

To my friends, who have been nothing but understanding when I have disappeared, but never stopped shouting from the rooftops whenever I shared some exciting news; I feel so lucky to have them.

A resounding thank you to the organisations, landowners, managers, and members of the public who I have worked with; I am so appreciative for their time and expertise, access to their land or data, and the insane collaboration opportunities that I could not have dreamed of.

I would like to thank my examiners and Viva convenor for the time taken to review my thesis and for their forthcoming feedback.

My research has been supported by a University of Glasgow's College of Social Sciences PhD Scholarship. I am so grateful to have had the opportunity to explore Scotland's natural beauty whilst endeavouring to conserve it. It has only made me more determined to allow people to experience all that the natural world has to offer. Most do not get the chance to pursue what they love, and I will never be able to express the emotion I feel when I say that I have had, and hope to continue having, this experience.

Author’s Declaration

I declare that I am the sole author of the work contained within this thesis, except where explicit reference is made to the contribution of others, and that it is of my own composition. No part of this work has been submitted for any other degree at the University of Glasgow or any other institution.

Samantha Suter

October 2023

Abbreviations

AGB	Above Ground Biomass
BAP	Biodiversity Action Plan
BSBI	Botanical Society of Britain and Ireland
CBD	Convention on Biological Diversity
CEH	Centre for Ecology and Hydrology
CS	Citizen Science
CV	Coefficient of Variation
DMP(s)	Data Management Plan(s)
EBV(s)	Essential Biodiversity Variable(s)
EUNIS	European Nature Information System
EVI	Enhanced Vegetation Index
GBIF	Global Biodiversity Information Facility
GEE	Google Earth Engine
GVI	Greenness Vegetation Index
HabMoS	Habitat Map of Scotland
LAI	Leaf Area Index
MS	Micasense
NDVI	Normalised Difference Vegetation Index
NGO	Non-governmental Organisation
NIR	Near-infrared
NVC	National Vegetation Classification
OS	Open Science
PLS	Partial Least Squares
PS	Planetscope
RF	Random Forest
RMSE	Root Mean Square Error
RS	Remote Sensing
S2	Sentinel-2
S2REP	Sentinel-2 Red-edge Position
SAR	Synthetic Aperture Radar
SD	Standard Deviation
SDG(s)	Sustainable Development Goal(s)

SMCP	SPAD-Measured Chlorophyll Proxy
SPAD	Soil Plant Analysis Development
SRG(s)	Species-rich grassland(s)
SVH	Spectral Variation Hypothesis
SVM	Support Vector Machine
SWIR	Short-wave Infrared
UAV	Unoccupied Areial Vehicle
UK BMS	UK Butterfly Monitoring Scheme
VI/VIs	Vegetation Index/Indices

Chapter 1. Introduction

1.1 The Global Biodiversity Crisis

The Earth is currently experiencing what is known by many as the Anthropocene; a human-mediated geological era influenced by human population increases, impacting the Earth's biodiversity and functioning (Ceballos *et al.*, 2017). Anthropogenic activities include habitat destruction, pollution, burning fossil fuels, and overfishing. As such, a range of environmental effects have ensued, from ocean dead zones, human-influenced climate change, and large areas of habitat loss (Malhi, 2017).

This biodiversity crisis has led multiple organisations and governments to create targets for mitigating and reversing these impacts. For example, the Convention on Biological Diversity (CBD) brought together 150 governments in the 1990s to develop a plan for sustainable living, resulting in the formation of the Aichi targets to be addressed between 2011 to 2020 (Convention on Biological Diversity, 2011). These 20 targets for tackling the biodiversity crisis focus on broader statements concerning biodiversity, including: "a) addressing the underlying causes of biodiversity loss by mainstreaming biodiversity across government and society; b) reducing the direct pressures on biodiversity and promoting sustainable use; c) improving the status of biodiversity by safeguarding ecosystems, species, and genetic diversity; d) enhancing the benefits to all from biodiversity and ecosystem services; and e) enhancing implementation through participatory planning, knowledge management, and capacity building" (Convention on Biological Diversity, 2020). However, it appears that these have yet to be met, with more recent assessments showing increasing declines across global biodiversity (IPBES, 2019). The IPBES (2019) assessment, which evaluates the current state of biodiversity, found that Aichi targets, such as awareness of biodiversity issues and how to conserve biodiversity, have only been partially met, whilst targets including halving habitat loss and decreasing habitat fragmentation (the biggest drivers of biodiversity loss) have been inadequate.

One aspect influencing our ability to achieve (or understand progress towards) these targets is the pressing need to monitor these rapid changes in biodiversity. Without such monitoring, it is impossible to identify population trends, causes of changes, and potentials for solutions. As such, biodiversity monitoring is needed to prevent ongoing species loss and allow continued ecosystem functioning to ultimately support humankind and biodiversity alike.

1.2 The Status of Traditional Biodiversity Monitoring

In the past, biodiversity monitoring has been applied in various contexts to reach a multitude of aims (Henry *et al.*, 2008). These aims may be to address scientific questions (such as the biological basis of specific phenomena) by researchers (Couvet *et al.*, 2011; Nichols and Williams, 2006); inform conservation or land management decisions by ecologists, landowners, and businesses (Lindenmayer *et al.*, 2012); be "curiosity-driven" by leisure groups or individuals (Pocock *et al.*, 2015); or provide data for human health and well-being (from inferences of water or air quality, for example) through biodiversity indicator analysis (Ansari *et al.*, 2017; Muhamedieva and Niyozmatova *et al.*, 2023). Whilst traditionally, biodiversity monitoring initiatives may have been standalone to reach individual goals, it is now more widely accepted that global biodiversity monitoring requires collaboration to become all encompassing (Anderson, 2018; Navarro *et al.,* 2017; Perino *et al.,* 2022).

These various aims could have been met through a multitude of biodiversity monitoring methods, of which there may be overlap across these aims. Methods traditionally, and still to this day, include i) certain counts and density measures based on definitive numbers (of a species population, for example) (Gaines, 1999), ii) indices or proxies that are determined from other factors (e.g. diversity in sound recording from acoustic devices (Alcocer *et al.,* 2022)), and iii) measured units or samples as a subset, or estimate, of biodiversity (Montes *et al.,* 2021).

In more recent years, biodiversity decline has drawn attention from multiple stakeholders to try combat the ecological crisis (Wallington *et al.,* 2005; Kühl *et al.,* 2020; Sterling *et al.,* 2017). This engagement can be two-fold; seeing value in what nature provides for humans versus seeing value in nature's individual identity (Rülke *et al.*, 2020). It would be wise to increase the recognition of the latter, as drawing on this increases pro-environmental behaviours (Taylor *et al.*, 2020*a*; Taylor *et al.*, 2020*b*). However, since the 1990s, human-orientated strategies are now adopted more readily, for example, with the establishment of the Earth Summit making monitoring of biodiversity law across nations or through the concept of 'Nature-based Solutions' (Boyle and Sayer, 1995; Seddon *et al.*, 2020). Protocols were adopted across countries to monitor biodiversity on local and regional scales, with strong emphasis on what biodiversity loss means for humans. Usually, biodiversity monitoring is governed by one overriding national organisation, split into smaller specialities each dealing with an independent area of biodiversity. However, this has led to duplications, imbalanced focuses, and unclear aims amongst departments (Camacho, 2020; Hagerman *et al.*, 2021; Lee *et al.*, 2005).

Prominent drawbacks in traditional biodiversity monitoring include the time required and associated costs of data collection. Large numbers of cumulative researcher hours and costly equipment are needed to successful quantify key species and habitats in monitoring programmes, and this has not been feasible globally (González-Oreja *et al.*, 2013; Kindsvater *et al.*, 2018). Data collection issues largely occur due to lack of finance, infrastructure (both for data collection and communication), and location inaccessibility (Amano *et al.,* 2016). As such, most monitoring programmes have focussed on proxies for species richness and diversity, such as the Simpson's Diversity Index, due to their simplicity and applicability. However, indices and metrics in biodiversity monitoring do not provide information on why these states have occurred e.g., due to the potential presence of invasive species, and are not all encompassing, disconnected from the information they are trying to represent (Lamb *et al.,* 2009; Marshall *et al.,* 2020). Many indices and measures are also based on estimates, leading to the potential to misrepresent species' population sizes because of detection error arising from the method employed, such as mark-and-recapture. To increase the scale of representation in monitoring programmes, measures are often extrapolated from small sample sites resulting in a survey error (Valdez *et al.*, 2023). Many programmes also focus highly on species that are either rare or have a greater role in ecosystems or society, meaning more common species are often overlooked until declines are much more severe (Hoye *et al.,* 2022; Oliver *et al,* 2021; Yoccoz *et al.,* 2001).

While the overarching goals of biodiversity monitoring are similar across nations, typically varying by taxonomy and drivers of biodiversity loss, there are in fact large gaps in the scope of biodiversity monitoring due to an array of barriers; in short these include poor study design and vague outcomes, lack of funding and collaborative support, and spatial, temporal, and taxonomic gaps (Lindenmayer *et al.,* 2012).

1.3 Gaps and Barriers in Traditional Biodiversity Monitoring

1.3.1 Taxonomic Bias

One of the biggest issues facing biodiversity monitoring is the scale and number of species or habitats that must be incorporated. It is found in many schemes that there are biases towards specific species or taxonomic groups. For example, monitoring data is lacking for 46% of IUCN (International Union for Conservation of Nature) listed fish species, but only 0.6% of listed bird species (Amano *et al.,* 2016). What is more is that only 1.7% of invertebrates (with intraclass biases seen, for example, coverage within Insecta is largely limited to Odonata) and 10% of plant species have been assessed (Hochkirch *et al.,* 2020). When looking at other biodiversity indices, such as the Global Biodiversity Information Facility (GBIF), insects were found to have the lowest representation (with a high intraclass bias as well) due to reasons including difficult identification, lack of interest, and a high proportion of undiscovered species – which is an estimated 80% (Stork, 2018; Troudet *et al.,* 2017). Perhaps unsurprisingly, increasing monitoring of insect taxa has been highlighted as vital for completing the Aichi targets and alternate methods are needed to reach this, as current practices will not manage this in the sought time frame (Girardello *et al.,* 2018).

It is difficult to identify what species should be involved in monitoring programmes, as each may have a vital function in an ecosystem. However, it has been suggested that indicator or priority species may be vital targets, as these can signify the health of an ecosystem that supports a high species richness. Although encompassing all species is vital, focusing on multiple species or ecosystems at once can lead to increasingly ambiguous aims, resulting in failed monitoring attempts (Pocock *et al.,* 2015). Therefore, targeting a specific habitat or set of species of which their monitoring can benefit the wider ecosystem and multiple species may be the best approach.

1.3.2 Spatial and Temporal Limitations

In addition, monitoring may also be highly spatially or temporally heterogeneous. Geographically, most monitoring programmes are based in the Northern Hemisphere and reaching countries in the Global South is an ongoing goal (Amano *et al.,* 2016). Many countries where biodiversity monitoring does not occur to a large extent, or at all, are areas where biodiversity is much higher – mainly in the tropics (Collen *et al.,* 2008; Hochkirch *et al.,* 2020). The geographical biases apparent in global biodiversity monitoring consequently result in a habitat bias as well. Even in the Global North, where monitoring occurs much more frequently, gaps across habitats are present, with temperate forests or habitats that exist inside protected areas being focused on (Martin *et al.,* 2012). When considering habitat monitoring, efforts must also concentrate on condition and not just extent (IPBES, 2019). Although geographical gaps are prominent in biodiversity monitoring, this is not to say that monitoring should completely focus away from areas that are currently intensely monitored and concentrate on data deficient regions entirely. Instead, monitoring programmes should consider how they can expand or be applicable in these geographical areas where data is lacking.

Temporal gaps in monitoring also occur for multiple reasons. It has been suggested that many monitoring projects usually operate for up to three years, whilst at least 10 years of monitoring is needed to accurately assess changes in biodiversity (Stephenson, 2020). Although there are now established long-term monitoring schemes (such as the UK Butterfly monitoring scheme), many gaps still occur which follow the same spatial and taxonomic biases (Ondei *et al.,* 2018). In a lot of areas baseline data is lacking (specifically for community assemblages), or inaccessible, resulting in unclear trends for multiple biodiversity targets across decades and since anthropogenic activities fuelled these changes (Hoye *et al.,* 2022; Magurran *et al.,* 2010). It is demonstrated that most monitoring programmes started after approximately 50% of the present impacts from human activities had already ensued (Mihoub *et al.,* 2017).

Geographical and temporal coverage is largely affected by cost and time restraints; however, language barriers and conflict may also affect monitoring (Amano and Sutherland, 2013). Long-term monitoring is particularly influenced by consistent funding, whereas frequent visitation rates may not be possible for certain natural processes or conditions. Other unforeseen circumstances may occur where monitoring cannot continue. For example, during the COVID-19 pandemic many in-person surveys had to stop operating, whilst some areas where monitoring is conducted closed access to the public. Conflict may also affect the long-term availability of certain monitoring schemes as it does largely across Africa where political instability is prevalent (Siddig, 2019).

1.3.3 Public Understanding and Policy Implementation

Comprehensive biodiversity monitoring also requires collaboration across stakeholder groups, from governmental bodies to the public. It is apparent there is a disconnect between biodiversity knowledge and the public. Reasons include misunderstanding of terms, such as “biodiversity”, disinterest in or inability to being outside, and failed media communication (Navarro-Perez and Tidball, 2012). This separation creates less agency for biodiversity conservation. Hooykaas *et al.* (2019) introduced the concept of species literacy, an understanding of native species as a representation of the term “biodiversity”. Low species identification scores were particularly prominent in children, with only an average of 35% of species successfully recognised (Hooykaas *et al.,* 2019). It can be argued that creating a broader species-based knowledge in the public will cultivate the need for monitoring a greater range of taxa.

Gaps still exist in the end use of biodiversity monitoring data with large data sets accumulating but not having any practical application (Ruckelshaus *et al.,* 2020). Lack of policy implementation following biodiversity data collection can result from reduced accessibility due to various restrictions (Geijzendorffer *et al.,* 2016). Additionally, due to the multi-field nature of policy legislation and conservation biology, coupling these disciplines is not straightforward (Rydén *et al.,* 2020). The absence of specific ecosystem service monitoring, to evidence how ecological data merges with societal functioning, makes creating relevant policies difficult (Navarro *et al.,* 2017).

Maes *et al.* (2012) identified gaps in the translation of biodiversity monitoring of ecosystem services, such as 1) medicinal resources, 2) the importance of genetic diversity for disease prevention, 3) keystone species for habitat functioning, and 4) that this monitoring must look outside of only water and food resources and climate regulating services, to ensure unbiased policy decisions. Previously collected biodiversity research can be used to create actionable measures through mobilising data with collaborative networks and improved accessibility of archived datasets. For areas where data has not been collected, it appears more beneficial to identify a biodiversity conservation policy goal and collect data that will inform this policy (Geijzendorffer *et al.,* 2016; Wetzel *et al.,* 2018).

1.4 How can Biodiversity Monitoring be Improved?

Paucity of biodiversity monitoring data, driven by the factors above, may be addressed by making monitoring more efficient and encompassing of all biodiversity. For example, directed programmes toward conservation targets that are outlined by organisations, governments, or international bodies (such as the CBD) (Kühl *et al.,* 2020). However, the biases in conservation policies need effective collaboration between scientists and governments to ensure these are addressed. In Europe, multiple conservation policies are in place (for example, Natura 2000, a network of protected areas across Europe, are managed under similar sustainable practices across EU member states (Evans, 2012)) but large overlaps of species coverage were found (mostly mammals and birds) (Henle *et al.,* 2013). Henle *et al.* (2013) suggested that an overarching system of biodiversity targeting should be developed first, prioritising species or communities based legal protections, use as indicators of wider ecosystem health, or providers of ecosystem services, allowing monitoring to be directed to where it may have the greatest benefit.

When creating policy for biodiversity monitoring, reducing gaps can also be realised through the creation of translatable goals or tools that can be used across taxa, habitats, and countries. A call for a “culture of integration” in biodiversity monitoring has been made by Kühl *et al.* (2020) where communication globally allows monitoring to be harmonised, enabling this complete capture of biodiversity data. This is realised through a global united monitoring network supported by active stakeholder engagement. Creating this network will be possible through various tools; utilising the benefits of each to help address gaps resulting from limitations of certain methods (Stephenson, 2020). What is clear is that an increase in communication and collaboration, coupled with a merging of disciplines, will allow bridges to be formed in the global fight against biodiversity decline. This can be done through the practice of Open Science.

1.5 Tools for Improved Biodiversity Monitoring: Open Science

Open Science (OS) is "an effort to close the gap between science and society by democratising scientific knowledge" (Holbrook, 2019). OS can be applied to all areas of the research process; from publishing under an open access agreement to including stakeholders and members of the public in scientific design (see Figure 1-1). In biodiversity monitoring research, OS can be instigated through 1) collaborative research design (engaging the public in the monitoring of their local areas, for example), 2) the use and open publication of data management plans (DMPs), preregistrations or registered reports, 3) utilising free software, and 4) openly publishing raw data and monitoring results (Bowman and Keene, 2018). To implement the use of OS in biodiversity monitoring, its applications and benefits must be explored.

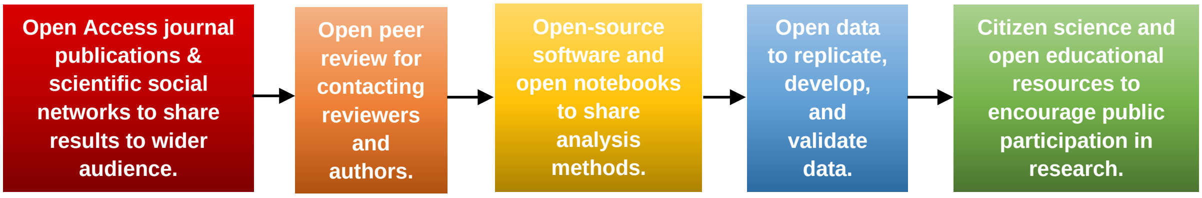


Figure 1-1. Open Science research practices increasing in openness from left to right. Figures adapted from Bowman and Keene (2018) and de la Fuente (2019).

Although the concept of OS can be traced back to the 17th century, it is really since the turn of the 21st century that OS has had wider recognition (Gong, 2022). It appears that OS has not been historically implemented largely in the various biodiversity monitoring approaches. Scientific research publications surrounding biodiversity conservation since 2000 found that most research (>95%) did not adhere to the true definition of open access publishing, for example (Fuller *et al.*, 2014). Even where other OS practices may be in use such as with collaborative research, the previous major issue of implementing OS in biodiversity monitoring initiatives seems to relate to the final sharing and dissemination of biodiversity information (Gaikwad and Chavan, 2006).

Information on how private companies, businesses, and recreational groups observe OS practices in biodiversity monitoring is little researched. It could be assumed that in past private biodiversity monitoring may have only shared findings where necessary, e.g., with partners, local councils, or governments. This is supported by the evidence that most biodiversity information appears to come from few organisations, which largely have a biodiversity conservation agenda (Stephenson and Stengel, 2020). On the other hand, recreational biodiversity monitoring in general may be more open due to the nature of public engagement. However, with a greater focus on national and international targets for addressing biodiversity loss, and with a rise in public concern regarding the ecological crisis, it is becoming more prominent and even required for monitoring bodies to, at the minimum, report their conclusions (Hassan *et al.,* 2022; Smith *et al.,* 2019). Most research on the use of OS across biodiversity monitoring appears to be focused on information sharing, and more implementation of OS practices must occur throughout the entire research process (Roche *et al.,* 2022).

1.5.1 The Pros and Cons of Open Science

Applications of OS are not without challenges. Confusion and opposing opinions shroud the movement, making its implementation less straightforward. When researchers considered OS to access data for reproducibility or papers for citing purposes, attitudes to OS were largely positive. In comparison, when these same researchers considered their own work being readily available this was met with caution (Nicholas *et al.,* 2019). Lack of clarity in OS is not limited to who moderates OS, the use of variable terminology (open access, open data etc.), and what is considered OS and what is not (Nicholas *et al.,* 2019). Some journals will state they are open access but only allow publishing under the most limiting Creative Commons license, whilst others will only enable open access for papers that are not yet published but only In press (MacCallum, 2007). Misunderstandings also occur as journals do not promote their OS policies resulting in a lack of awareness, whilst some authors state the requirements or software to enable OS are simply not there (Van Noorden, 2014).

Along the research process, OS obstacles may exist. Williams *et al.* (2017) found that a common issue with DMP requests is inconsistencies in the requirements. Preregisters and registered reports are often thought to limit researcher freedom, be more time consuming, and risk plagiarism (Klein *et al.,* 2018; Sarafoglou *et al.,* 2022). Consequently, DMPs and preregisters were found to be amongst the lowest of implemented OS practices in biodiversity citizen science projects (Suter *et al.,* 2023) (see chapter 3). To address these concerns, frameworks and training should be, and are being, created (Williams *et al.,* 2017). The aim of OS practices is to increase transparency, not to reduce flexibility, as these processes allow the visualisation of errors and prove to rectify themselves in terms of increased data quality (Nosek *et al.,* 2019).

When deciding where to publish research findings it should be encouraged to use open access journals, such as the Public Library of Science to university publishers like UCL Press. A greater number of citations can be generated by open access publishing, which is evident across multiple disciplines (Clements, 2017; Hajjem *et al.,* 2006). A higher number of citations can also increase the potential of coverage on social media platforms or newspapers. This has been demonstrated with the use of Twitter, whereby citations rose in ecology research because of sharing via the social media platform (Finch *et al.,* 2017; Lamb *et al.,* 2018). OS has also created new software that improves research practice. Software includes websites for storing, sharing, and analysing data, such as Enlighten or R Toolbox which investigates discrepancies in correlation which values in reports (Allen and Mehler, 2019). Having access to raw data and code increases collaboration by combining expertise across multiple fields and can also be used for educational purposes, allowing research methods to be understood on a practical level (Saluja and Thakur, 2020). Although there is software available for OS, many are still in the process of being created and understood.

For OS to be successful, there must be existing tools facilitating the availability of research and standard guidelines in place to avoid confusion or the potential to take advantage of easily accessible data. These necessary tools for OS are mainly in the form of online repositories and databases to store data, social media platforms for sharing data, and free access journals to present research outcomes (Neylon and Wu, 2009). Universal OS practices may not apply to all stages of research and data which require different tools. It may be more advantageous to actively encourage OS through policies, institutions, and funding bodies, whilst allowing the researcher to justify where OS applies to them and having their work recognised through strict credit checks (Levin *et al.,* 2016). Cultural change towards practicing OS is often noted as the most difficult task when trying to make OS the norm. OS workshops focusing on approaches to OS and why it should be practiced should be made available where applicable, to breed an understanding of its importance (Ignat and Ayris, 2020). Hopefully, as OS becomes more conventional, tools to facilitate it will develop and the practice will become more standardised, clarifying the process.

1.6 Tools for Improved Biodiversity Monitoring: Citizen Science

Another widely used tool for biodiversity monitoring is citizen science (CS). Simply put, CS is the involvement of “the general public in scientific research tasks” (Vohland *et al.,* 2021). Participant involvement can make up different types of CS. Contributory projects involve volunteers who are usually only associated with data collection; collaborative projects include volunteers associated with data collection and other aspects of the project; and finally co-created projects have volunteer participation throughout the entire research process (Bonney *et al.,* 2009). Due to CS’ nature of involving the public in data collection, and more recently in project design, it is at its core an OS approach and facilitator.

CS is not a new concept and the rise in CS is well documented (Berti Suman and Alblas, 2023; Kosmala *et al.,* 2016; McKinley *et al.,* 2015). In traditional biodiversity monitoring, CS will have had its largest contribution (although not exclusively) in recreational applications, as it was predominately created for such purposes. CS applications have previously been evidenced in ecological small scientific research projects, leisure pursuits, natural history collections, and even national programmes (such as for water monitoring) (Miller-rushing *et al.,* 2020). CS may have previously been utilised least in monitoring by private companies and businesses, as these largely would have hired internal staff or contracted professionals. However, where CS examples are more readily utilised within NGOs, for example, there is now even more scope and potential for all organisations that undertake biodiversity monitoring to engage with CS in some way (Anderson *et al.,* 2020). Even without creating their own CS projects, businesses are being more heavily encouraged to utilise crowd-sourced data (Cappa, 2022).

With further use and recognition of CS, there is no doubt it will be continuously applied in biodiversity monitoring.

1.6.1 Citizen Science and its Contribution to Open Science

Research shows that CS works well to entice public participation where the interests of the public are considered and met from the start of the project; where communication about the project reaches the target audience over a variety of means (e.g., social media and local newspapers); and where incentive is given, for example, through recognition or competitions (Dickinson *et al.,* 2012). Usually, participation in CS comes from an audience who is already partially interested in the specific research of a project or science in general, and this needs to be considered when designing CS projects and how it impacts on OS (Land-Zandstra *et al.,* 2021; Pers. obvs., 2022). This suggests that an interest in science must be bred early on in life, possibly through curriculum-based learning during the school years, to create a larger potential source of participants, and highlights the importance of targeting the audience in the advertisement and design of projects (Martin, 2017). A focus when designing CS projects, or considering OS in general, should be how to enlist interest from members of the public who previously have had little interest in scientific research.

If a goal of OS is to democratise science through an increased public understanding of science, and interest in governance, this is where CS as a tool seems to fall short, and increased knowledge of a research area has not yet presented itself in the form of social change. Evidence suggests that when CS projects have the intention to inform policy, it is possible that governance can be influenced by the results of the project, and participants are more likely to be involved in the process (Hollow *et al.,* 2015; Warner *et al.,* 2019). What needs to be considered in the design of a CS project is what continual impact and involvement can participants have past the ending of the project, such as through contributing to local councils, environmental stewardship, and informing policy (Toomey and Domroese, 2013).

If clear guidelines around data sharing are set out, standardised, and monitored, these should address any concerns around sharing CS data. Where data is open and available the combination of datasets has led to large collaborative studies, specifically targeted to climate change, and changing trends of invertebrate populations (Follet and Strezov, 2015). If CS data becomes more open, this should lead to increased collaboration between researchers and project designers and benefit the OS movement in many ways.

1.6.2 The Importance of, and Concerns in, Citizen Science

One main draw for CS is its potential to collect a vast amount of data over increased space and time, which would not be possible with a smaller team of researchers, or require large funding (Dickinson *et al.,* 2010). This high participation in a project will also increase efficiency due to greater data collection capacity, and results can be reached quicker. CS, like OS, however, is not only for the scientist. It benefits the public that is involved in the projects through increasing knowledge in a particular area, building unique and transferable skills, creating a community to be a part of, and allowing the public to have a say in informing policy (from increased engagement and awareness of scientific issues) (Adler *et al.,* 2020). This has the potential for social change and impact decision making, as the public is more involved with an issue or area of research (Butkevičienė *et al.,* 2021). The public can also benefit the research by bringing in their own local or outside knowledge through participatory research, much in the same way that collaboration in OS can combine the expertise of researchers across multiple fields (Tengö *et al.,* 2021). CS projects may also be cost-effective due to the designs of the project, often using apps or websites for data collection, with participation being voluntary. This could mean that CS is applicable in developing countries where expensive data collection and equipment may not be feasible (Schröter *et al.,* 2017).

A large concern when considering CS as a valid form of scientific research is the competency of the public conducting the research (Balázs *et al.,* 2021). Participant ability appears to vary depending on the required data collection methods, demographics of the observer, and previous experience, but can be improved with detailed protocols, personalised training, pre-sampling tests, and continuous practice (Dickinson *et al.,* 2010). It may also be possible to account for any variations in data collection amongst participants in the analysis stages, if participation and data collection are large enough (Theobald *et al.,* 2015).

Retaining participation is also key for long term programmes, as the amount of data depends largely on the level of participation of the public (Schröter *et al.,* 2017). As participants are usually unpaid, other forms of incentive should be considered. This can be in the way of skill-building, work experience, connecting with experts of that field, or some form of other recognition, for example, certificates for the number of hours worked/reaching data collection targets. It is also a worry in CS that research is based on incentive rather than being hypothesis-driven, and, as such, is not “value-free” (Elliot and Rosenberg, 2019). This essentially means that the drive behind the research is not to do with truth-finding, but is agenda ridden. This may not necessarily be the case, but even if it were, this need not be a bad thing. The issue here is not of value, as by definition research has value, but rather bias. It is possible that bias can impact on CS research, either by putting effort into one area of research over another, or in the analysis. However, with CS forming a part of OS, bias can be reduced or identified due to the greater transparency of the work and, thus, higher levels of verification.

Although CS is not a new concept, its uptake on a large scale has only increased rapidly over the 21st century (Kullenberg and Kasperowski, 2016). As such, there are still some limitations across CS and areas where improvements can be made. For example, most CS projects are established in the Global North (Pocock *et al.,* 2019). CS can, and should, be expanded globally across all cultures and geographic areas. However, if CS is to be undertaken in developing countries, it must be ascertained that local populations are not taken advantage of, and projects must ensure a strict working relationship for the project to be successful.

Theobald *et al.* (2015) found that most of the research from CS projects are not published in peer-reviewed journals, irrespective of the quality of the data collected. This may be due to the sensitive nature of CS data that is collected, for example, the location of protected species, or personal information that cannot be anonymised. Therefore, the research often does not reach the masses, contradicting a focal purpose of CS. A hindrance to the potential of CS may also largely be found in the attitude of traditional scientists and researchers, as the data collected may not be considered genuine (Theobald *et al.,* 2015). However, with the increasing creation of CS projects, that not only collect valuable data, but expand the interest and knowledge of the public and create social change, the acceptance of CS as a valid form of scientific research will follow.

1.6.3 Best Practices and Improvements in Citizen Science

Although CS is growing in its applications, the practice does call for some standard conditions; recruitment, retention, and incentive of the participants, methods that are replicable and can be followed easily and individually, and data that is verified and trustworthy (Worthington *et al.,* 2012). Wiggins and Crowston (2012) created a typology of CS to categorise different projects based on the project aims. The framework simplifies project types and identifies issues that may be faced. As there is likely to be an overlap between categories, and due to the nature of CS, potential problems still fall into common themes; issues with data whether that be from sampling bias versus volunteer experience, or sustainability issues either from funding or volunteer participation (Adler *et al.,* 2020; Wiggins and Crowston, 2012).

There are multiple ways that data quality can be assured in CS, and various methods can be applied to a project to confirm the highest data quality possible (Table 1-1). The level of expertise needed by a participant is likely to influence the quality of the data that is collected; if greater experience is required for a project, the likelihood of errors in data collection increases, for example, misidentifying species. Approaches to tackle this could be with highly detailed protocols available to participants, implementing a buddy system whereby an expert is paired with someone less experienced, or creating a self-confessed experience criterion that can be used for targeted training approaches (Crall *et al.,* 2011). Equipment and ease of the project also contribute to the quality of the data, with pilot studies being useful to identify if the procedures of the project are simple enough to follow, especially for an unspecialised audience (Kosmala *et al.,* 2016).

Table 1-1. Summary of data quality issues identified in citizen science projects, methods to address the issues, and examples of projects and studies that have used certain techniques.

Issue	Solutions	Examples
Lack of experience of participants	In person training, Detailed protocols, Pilot study for methods, Buddy system	UKBMS (2019) Urban air quality citizen science (Cowie *et al.*, 2014)
Trust of users	Trust metrics	CoralWatch (Hunter *et al.*, 2013)
Data validation	Automated data entry forms, Validation by experts, Replication through participants, Statistical Analysis	HEES in BirdTrack (Wessels *et al.*, 2019)/Smart Filter in Project Feeder Watch (Bonter and Cooper, 2012) iNaturalist (Balázs *et al.*, 2021) Snap Shot Serengeti (Swanson *et al.*, 2016) Covariates in eBird (Johnston *et al.*, 2020)

CS requires volunteer participants to be established. When questioning participants for the reasons of their involvement, responses range from initial interest in the project topic or its outcomes, self-improvement either through skill-building or networking, social aspects, such as becoming a part of a community, or simply wanting to help (Rotman *et al.*, 2012). It is important to note the timing of motivation as well, as participant involvement is likely to decrease at certain stages of a project (Rotman *et al.*, 2012). Therefore, when recruiting and retaining participants, their varied motivational factors must be considered when designing a project to obtain the greatest number of volunteers. Retainment, for example, could be an issue of method simplicity, volunteer advancement, or lack of project outcomes (Domroese and Johnson, 2017).

Initial volunteer involvement in project design is important, but difficulties lie around public understanding of the scientific process. Most CS projects are contributory; Smith *et al.* (2017) found over 50% of 173 CS projects fell into this category. However, the public can be more involved by informing study design in areas including ethical considerations (e.g., consent forms), timelines of the projects, evaluating data collection techniques, and contemplating research outcomes. Methods for the inclusion of the public could involve focus groups, interviews, and questionnaires, but focus groups appear preferential due to gathering the most information in a reasonable timescale (Boote *et al.*, 2010). It has been recommended that an ongoing exchange between scientists and participants must exist across multiple platforms, not limited to social media engagement, social and educational events, constant sharing of results, feedback between all contributors, and understanding the needs of all volunteers (de Sherbinin *et al.*, 2021; Druschke and Seltzer, 2012).

Both data quality, project scope, and recruitment of participants in CS have been improved by technological advancements (Davies *et al.*, 2016). With these developments CS can be practiced in multiple ways, including on mobile phone apps, such as iNaturalist, or web-based projects, such as Galaxy Zoo (Nugent, 2018; Raddick *et al.*, 2019). The use of different platforms has expanded CS because it is able to reach a greater audience with varying interests. As technology has progressed, simplicity of digital CS platforms has improved, allowing greater participation in projects and data collection where traditional CS projects are not applicable (Jones *et al.,* 2018). The creation and use of autonomous sensors in CS have enabled further reach and depth of data that is collected and will only continue to do so as technology improves and becomes less expensive (Newman *et al.,* 2012).

A relatively new concept in CS is the use of RS data (satellite and aerial imagery of the Earth) for CS projects (Geller *et al.,* 2017). The use of citizens in conducting Earth observations has only increased in the last decade, where citizens can provide *in situ* observations to corroborate remotely sensed data (Fritz *et al.,* 2017). Remotely sensed data provides a large collection of data that can be analysed and used effectively across multiple disciplines to target potential research questions. One area, for example, is for identifying habitat and monitoring change, which is extremely relevant in today's ever-changing world. Very few CS projects have used RS data in the past, with most developed in the past five or so years (e.g., Heritage Quest to identify archaeological objects) (Lambers *et al.,* 2019). Therefore, there is a gap in the potential of both CS and RS that can be targeted through the combination of these methods.

1.7 Tools for Improved Biodiversity Monitoring: Remote Sensing

Existing methods within remote sensing (RS) include the generation of "information about objects or areas at the Earth's surface without being in direct contact with the object or area", using platforms, such as satellites and aircraft (Aggarwal, 2004). RS can acquire data on surfaces or areas of interest at multiple spectral wavelengths, going beyond the natural range of human vision (0.4 – 0.7 μm) and can be derived from a variety of sources, actively or passively. Passive sensors detect radiation that is reflected or naturally emitted from the Earth's surface, while active sensors, such as Synthetic Aperture Radar (SAR) produce their own source of illumination and detect the amount of radiation that is returned to the sensor (Campbell and Wynne, 2011). These different RS systems can capture data over various spectral bands, spatial resolutions, or temporal scales, each having their own advantages for various applications (Vali *et al.,* 2020). Choosing the correct sensor is heavily dependent on the intended application.

RS does not have a long-standing history in biodiversity applications as OS and CS do, due to its technological requirements. However, it has been ever increasingly employed in a variety of contexts over the last few decades: applications have focussed on natural resource or land use change monitoring, rather than specifically on biodiversity (Roy *et al.,* 2017). Conversely to OS and CS, RS has larger applications in the private sector of monitoring due to the costly nature of RS devices and some data. This has made it less accessible for NGOs or recreational biodiversity monitoring initiatives. However, there has been a rise in open access RS data that is allowing this to change, which is being more readily seen (Turner *et al.,* 2015). Although the use of RS techniques in scientific biodiversity monitoring research has risen in the last few years, there is still much greater potential for this to be achieved (Reddy, 2021; Reddy *et al.,* 2021). The increase in partnerships across public-policy-private relations, as well as open RS software and data, will allow RS to become more prevalent in biodiversity monitoring, answering the continuous calls for this application (Antonelli *et al.,* 2023; Khorram *et al.,* 2016).

1.7.1 Remote Sensing and its Applications for Monitoring

RS has one of its largest applications in global change research, for example, with the use of monitoring land-use cover (Asokan and Anitha, 2019). As technology has advanced, the application of RS has increased, making this an ever more critical resource for land-change monitoring. This is because RS can capture data over large spatial scales and incorporate areas that are otherwise inaccessible to humans. More frequently, RS has been applied in biodiversity conservation research (Petrou *et al.,* 2015). To be able to monitor these changes in land cover and use, the land must be classified into habitats depending on their unique features. There are certain habitat indicators, such as vegetation structure and biomass, that are vital for determining these land classifications (Klemas, 2001). These indicators are usually measured at ground level; however, it is a time-consuming process to collect this information across multiple landscapes (Wood *et al.,* 2012). Therefore, RS may be used to capture this data at a fine resolution but across a large spatial area.

RS has been employed largely across forest ecosystems, with other habitats, such as grasslands, more frequently overlooked (Ichter *et al.,* 2014; Reinermann *et al.,* 2020). This is due to greater difficulty differentiating grassland species composition due to their smaller spatial resolution (Zlinszky *et al.,* 2015). Grassland classification is also more challenging due to the greater intra-class heterogeneity within the habitat creating higher-level classifications, such as species-rich neutral versus calcareous grasslands. Grasslands also undergo multiple inter-annual changes from various flowering and senescence timings, indicating that sampling time can alter grassland classification too (Reinermann *et al.,* 2020). As such, there is a large gap in knowledge for grassland classification on a global scale, which is problematic due to the ecosystem services that these habitats can provide (Zlinszky *et al.,* 2014).

Technological advances and increased availability of RS data have shown it is possible to identify grassland plant communities from satellites, such as Sentinel-2 (depending on the size of the community), especially when creating a time-series of images, which reduces the effects of cloud cover (Rapinel *et al.,* 2019). However, plant species scale and phenological status may be limiting factors (Thornley *et al.,* 2022). Studies also suggest that Sentinel-2 data can accurately measure leaf area index (LAI) and above ground biomass (AGB), important in the identification of grassland habitats (Schwieder *et al.,* 2020). It must still be evaluated whether Sentinel-2 has a high enough spatial and spectral resolution for all indicators of grassland habitats compared to other RS sensors.

RS devices that can capture multiple wavelengths and have higher spatial resolutions are more beneficial for grassland identification, as they can classify characteristics of vegetation with greater accuracy. However, in certain circumstances and depending on the habitat type, satellites with a higher spectral resolution but lower spatial resolution, such as Landsat, outperformed satellites with a higher spatial resolution but smaller spectral resolution, such as SPOT (Nagendra *et al.,* 2013). Most grassland classifications have come from optical RS (Reinermann *et al.,* 2020). It can be questioned why the use of SAR is not more highly used, as SAR devices are not affected by weather conditions, removing any hindrances from cloud cover, which is advantageous over land areas that have high cloud cover for most of the year (i.e. the UK) (Dusseux *et al.,* 2014). SAR provides information at a single wavelength while optical or multispectral sensors can provide information at multiple wavelengths (Sommervald *et al.,* 2023) (Figure 1-2). When comparing Sentinel-1 imagery (SAR) to Sentinel-2 imagery (optical), Sentinel-1 was less competent at differentiating between habitat classifications that were similar in texture. This may pose a problem for distinction between discrete grassland classifications. There is potential for the combination of SAR and optical imagery to provide high accuracy and avoid weather related issues (Dusseux *et al.,* 2014; Meneghini, 2019).

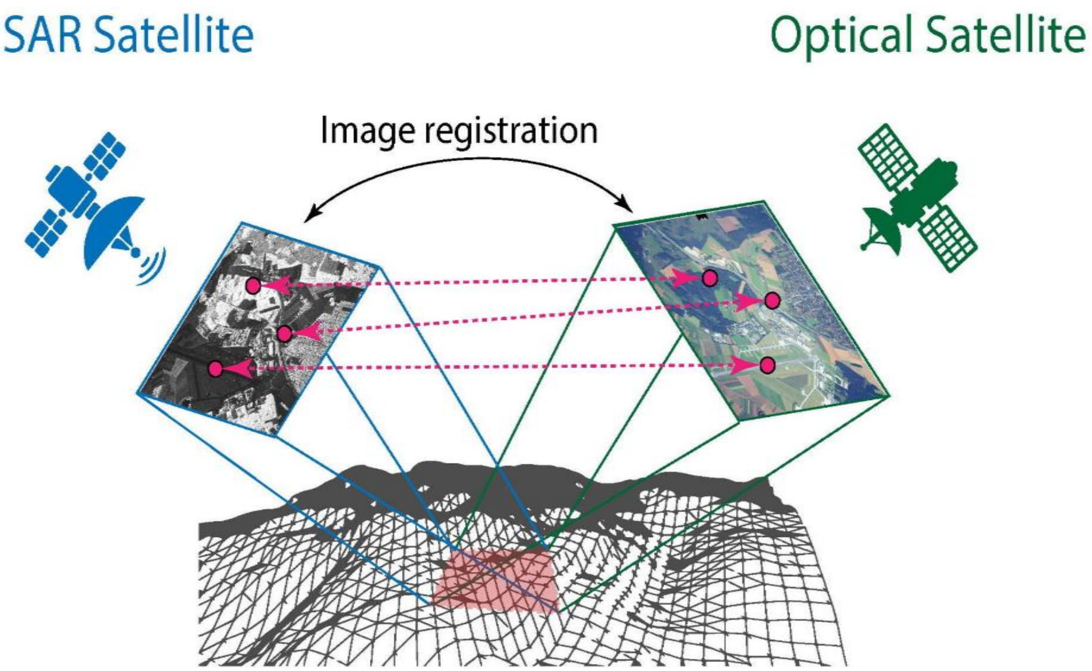


Figure 1-2. The different mechanisms of SAR (an active sensor using backscattering from microwaves to produce a single wavelength black and white image) versus optical (a passive sensor using reflected light from the sun to produce colour images from multiple wavelengths) satellite imagery. Image source: Sommervald *et al.*, 2023.

Light Detection and Ranging (LiDAR) imagery usually collects information at a single wavelength (across the near infrared for vegetation due to the backscatter of chlorophyll) (Dubayah and Drake, 2000; Lewis and Hancock, 2007). LiDAR can display vegetation height at a small spatial scale and may be an alternative method to satellite imagery, however, it is costly and not data abundant. Some limitations of its use are in windy circumstances; however, vegetation height can still be inferred, whilst other structures may not be. Although LiDAR is expensive, its increasing use is reducing overall costs and enabling easier access to the data. Comparing an unoccupied aerial vehicle (UAV), a pole camera structure from a Structure from Motion (SfM) photogrammetry (creation of a 3D structure based on a series of overlain images), and a hand-held LiDAR device showed that the LiDAR device was able to measure vegetation height most accurately in a grassland habitat (Obanawa *et al.*, 2020; Ota *et al.*, 2015).

Other methods of RS include the use of airborne devices. UAVs may be a preferable tool to capture images on grassland species composition compared to satellite imagery or higher airborne devices (as they can be used below cloud cover and reach inaccessible areas such as steep slopes) (Avtar and Watanabe, 2020; Quaye-Ballard *et al.*, 2020). This is truer for grasslands than for other habitats. On top of this, satellite imagery can have limitations including infrequent revisiting rates reducing the occurrence of data capture. UAVs can target this, as they can be used repetitively and take into consideration various seasonal differences in the appearance of certain species.

It seems the best possible RS tool to measure all indicators of habitat classes are those with both high spatial and spectral resolutions. The combination of RS methods may be beneficial to utilise the advantages of multiple techniques. Using satellite imagery complemented by aerial/UAV imagery or LiDAR can also minimise the impact of weather conditions interfering with the clarity of images (Wachendorf *et al.*, 2018). Finding these most advantageous combinations will allow greater overall accuracy of land classifications, from utilising the high accuracy of ground measurements to the high spatial reach of satellite imagery (Figure 1-3).

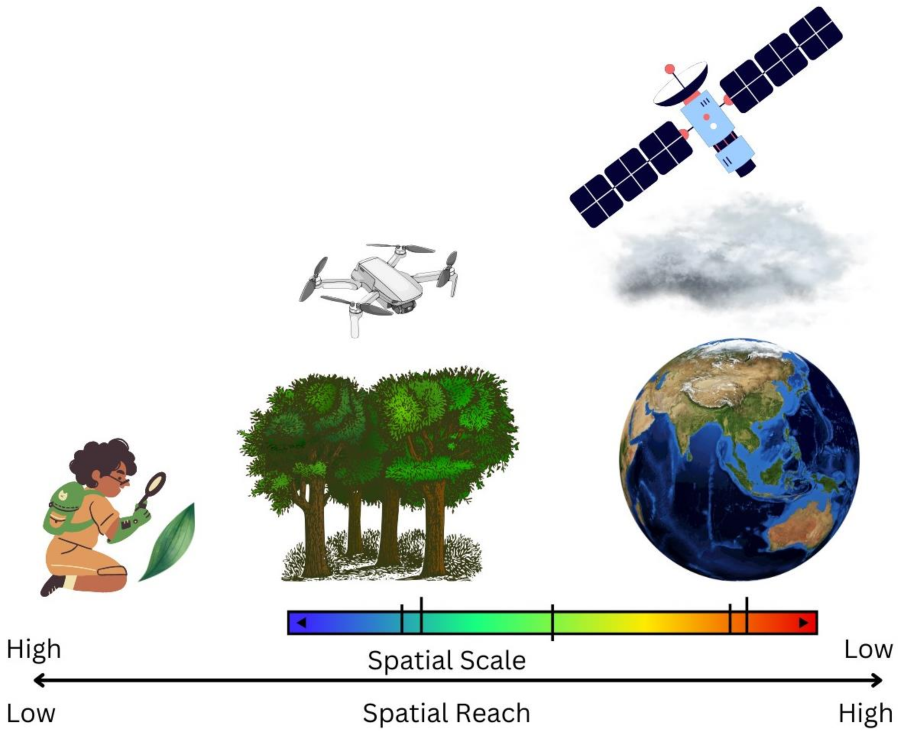


Figure 1-3. Comparison of spatial reach and spatial scale across remote sensing methods (UAV and satellite imagery) with *in situ* measurements. Figure adapted from: Bandopadhyay *et al.*, 2020 and generated in Canva.

In biodiversity monitoring, RS has been suggested as a tool to cover large spatial areas and be useful for increasing sampling frequency (Luque *et al.*, 2018). However, the two disciplines (RS and CS) have not yet combined to the extent that they can, and should be, due to lack of understanding and communication of needs between fields; conservation scientists are unsure of the quality of data that materialises from RS techniques or if these methods are more expensive than conducting biodiversity surveys (Borre *et al.,* 2011). There is a growing recognition that RS and CS can be combined to enhance these fields of knowledge, and a willingness to do so, but addressing the identified challenges will be paramount (Pettorelli *et al.,* 2014).

1.8 Combining Tools for Biodiversity Monitoring

Combining biodiversity monitoring tools (such as CS, RS, and OS) can allow synergistic efforts to benefit each other. Both CS and RS can address concerns over specific biodiversity gaps; whilst CS can increase frequency of monitoring to a certain extent, the use of RS can further enhance this. These tools can do this, as CS can utilise high participation to collect more data, whilst RS can target inaccessible regions or cover areas over a larger scale. The integration of techniques can address the challenges of each tool as well, for example, RS imagery can take a large amount of effort and time to classify, whilst CS projects may be limited spatially (Chandler *et al.,* 2017). RS data has doubled annually due to technological advances, which is beyond what is manageable by RS users. As a result, there is a lag in the assessment of this valuable data (Stephenson *et al.,* 2017). Uniting the public with RS scientists to utilise improvements in technology has great potential that is currently not being achieved. It appears that once the uptake of RS for biodiversity monitoring (for example, with the integration into a CS project) occurs, this could create a snowball effect to continue expansion in this area and maximise the benefits both tools bring.

A way in which the merging of these tools can be facilitated is through OS, to allow freely available RS data. Through collective OS tools, such as the Group on Earth Observations Biodiversity Observation Network and the GBIF, knowledge sharing is enabled for increased united research (Heberling *et al.,* 2021). As science becomes more open, the integration of methods across disciplines should allow greater scope for research to be undertaken, to target common goals. Cross-disciplinary approaches facilitated by OS can increase collaboration globally and incorporate data at an even larger scale. In the case of situations, such as the COVID-19 pandemic, the use of RS is even more applicable, as it will allow monitoring to continue whilst in-person surveying cannot occur (Sugai, 2020).

Calls continue for there to be national to global scale biodiversity monitoring networks, that will allow open data from a variety of monitoring tools to be integrated and enhance biodiversity conservation globally (Navarro *et al.,* 2017). Cord *et al.* (2017) questioned "*How can Big Data from CS and social media together with Earth observation be used to assess and monitor ecosystem services? Which conceptual and technical barriers must be overcome?*". Therefore, this collaboration will be essential for addressing the importance of monitoring to the public and policymakers (Stephenson *et al.,* 2017).

The current gaps restricting the integration of these methods largely focus on the technical demands of RS applications and the availability of collaborative platforms, software, and educational resources, along with long-held perceptions attached to CS data (concerns of data quality, mistrust, and traditional, often closed, mindsets) (Mazumdar *et al.,* 2017). However, it is not impossible to address these gaps. Technologies, especially those that are open, are increasing with the rise in open-source Copernicus satellite data, or 'Citizen Observatories', for example, that enable discipline integration to occur (Grainger, 2017; Viqueira *et al.*, 2020). Whilst integrative projects that demonstrate success allow preconceptions to be broken, accompanied by providing methods that foster trust (through training examples and data quality assurances) (Fritz *et al.*, 2017; See *et al.*, 2022).

It is possible to correct the gaps in biodiversity monitoring with targeted approaches, where a clearly defined aim can be reached using these tools of CS, RS, and OS. For example, indicators within five of the UN's Sustainable Development Goals (SGDs) had contributions from a range of CS projects, with the potential of its contribution to be much greater than this (Fraisl *et al.*, 2020). RS techniques could also be used to achieve Aichi targets that were missed by the 2020 deadline (O'Connor *et al.*, 2015). As these technologies continue to advance, they should be more heavily relied upon for integration into biodiversity targets.

1.9 Aim and Objectives

The tools of OS, CS, and RS have been analysed in their application of biodiversity monitoring to assess where each has its merits but also where improvements can occur, specifically through the merging of their uses. It has also been recognised that there is space for an instrument to be developed that utilises RS and CS for biodiversity monitoring as an OS approach. Therefore, this research aimed to 1) determine the role of OS approaches in the conservation of biodiversity, and 2) combine two approaches in biodiversity monitoring, RS and CS, to create an open and interdisciplinary habitat monitoring tool. This is with the hopes of providing a global method to benefit biodiversity conservation, whilst engaging multiple stakeholders to improve scientific democratisation. The thesis research questions may, therefore, help to address the concerns raised throughout this chapter for discipline merging and the importance of implementing OS.

As such, a specific research question was designed for each subsequent thesis chapter to meet the overarching aims of the thesis:

RQ1 – In which UK habitat may open science approaches and interdisciplinarity be employed to monitor biodiversity change? (Chapter 2)

A priority habitat is needed to act as a target of this thesis' research that an OS habitat monitoring tool can be applied to. The resulting case-study will help to address some of the gaps this chapter highlighted in deficiencies of biodiversity monitoring, focusing on taxonomic and habitat biases.

RQ2 – To what extent do citizen science studies of biodiversity demonstrate the principles of open science? (Chapter 3)

The research here questions if OS practices are utilised in biodiversity conservation applications, for example, in CS (an OS tool), and, if so, where are they employed and how can improvements be made throughout the research process? This will help demonstrate where OS can be utilised in the creation of biodiversity monitoring tools, especially interdisciplinary ones, to foster trust, and data and knowledge sharing. This is specific to the CS component as the resulting combined tool was a CS survey that utilised RS data. This research will not only guide the creation of my own survey but may help future research by identifying paucities in OS implementation, hopefully addressing concerns in the employment of OS in the scientific process.

RQ3 – Can a nationally relevant open-source habitat classification model be created for a priority habitat? (Chapter 4)

To effectively record habitats, the model needs to accurately identify them. If these habitats can be predicted reliably, this may help locate specific habitat for vulnerable species to reach national conservation targets. Therefore, the characteristics of this habitat and requirements for vulnerable species needs to be outlined. This research will provide the RS outputs specific to the biodiversity case study (outlined in chapter 2) that would be used in the CS survey. This information could also further inform stakeholders for national targets of habitat mapping and conservation efforts, where few classification models are applied outside of a study area.

RQ4 – Is currently available open-source remote sensing data able to accurately monitor priority habitats? (Chapter 5)

This research will look at other remote sensing applications for priority habitat monitoring, investigating whether there are open-source RS data at the spatial and spectral scale to improve monitoring attempts. This information will help address the gaps and concerns of creating the interdisciplinary tool by analysing if open-source remote sensing software, data, and platforms can help reach specified biodiversity targets.

RQ5 – Can an interdisciplinary, open science citizen science project be created, utilising remote sensing outputs for habitat monitoring? (Chapter 6)

This will ultimately align CS methods and RS data in the one open access tool for habitat monitoring. This research will provide insights into the process of creating such a tool, identifying if the open-source applications are available for this to occur, the technical requirements of the interdisciplinarity, and if knowledge gaps of cross-discipline research can be amended.

Each chapter also includes explicit research questions or objectives that are specific to their respective chapter research question, outlined above. The chapters specific questions are found at the end of each of their introduction sections. Therefore, this research aimed to 1) combine remote sensing and citizen science in a biodiversity monitoring tool, in an OS approach, 2) use the tool to help map species-rich grasslands across the UK and identify possible locations for the vulnerable Northern Brown Argus Butterfly, and 3) ensure stakeholder engagement, public interaction, and open science methodologies throughout the research for increased scientific democratisation. This is with the hopes of providing a global method to benefit biodiversity conservation, whilst engaging multiple stakeholders to improve scientific democratisation.

References

Adler, F.R., Green, A.M. and Şekercioğlu, Ç.H., 2020. Citizen science in ecology: a place for humans in nature. *Annals of the New York Academy of Sciences*.

Aggarwal., S., 2004. Principles of remote sensing. *Satellite remote sensing and GIS applications in agricultural meteorology*, pp.23-38.

Alcocer, I., Lima, H., Sugai, L.S.M. and Llusia, D., 2022. Acoustic indices as proxies for biodiversity: a meta-analysis. *Biological Reviews*, *97*(6), pp.2209-2236.

Allen, C. And Mehler, D.M., 2019. Open science challenges, benefits and tips in early career and beyond. *PloS biology*, *17*(5), p.e3000246.

Amano, T. and Sutherland, W.J., 2013. Four barriers to the global understanding of biodiversity conservation: wealth, language, geographical location and security. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1756), p.20122649.

Amano, T., Lamming, J.D. and Sutherland, W.J., 2016. Spatial gaps in global biodiversity information and the role of citizen science. *Bioscience*, *66*(5), pp.393-400.

Anderson, C.B., 2018. Biodiversity monitoring, earth observations and the ecology of scale. *Ecology letters*, *21*(10), pp.1572-1585.

Anderson, A.A., Williams, E., Long, M., Carter, E. and Volckens, J., 2020. Organizationally based citizen science: Considerations for implementation. *Journal of Science Communication*, *19*(3), p.A01.

Ansari, A.A., Shalini Saggu, S.S., Al-Ghanim, S.M., Abbas, Z.K., Gill, S.S., Khan, F.A., Dar, M.I., Naikoo, M.I. and Khan, A.A., 2017. Aquatic plant biodiversity: a biological indicator for the monitoring and assessment of water quality. In *Plant biodiversity: monitoring, assessment and conservation* (pp. 218-227). Wallingford UK: CABI.

Antonelli, A., Dhanjal-Adams, K.L. and Silvestro, D., 2023. Integrating machine learning, remote sensing and citizen science to create an early warning system for biodiversity. *Plants, people, planet*, *5*(3), pp.307-316.

Asokan, A. and Anitha, J., 2019. Change detection techniques for remote sensing applications: a survey. *Earth Science Informatics*, *12*(2), pp.143-160.

Avtar, R. and Watanabe, T. eds., 2020. *Unmanned Aerial Vehicle: Applications in Agriculture and Environment*. Springer.

Balázs, B., Mooney, P., Nováková, E., Bastin, L. and Arsanjani, J.J., 2021. Data Quality in Citizen Science. *The Science of Citizen Science*, p.139.

Bandopadhyay, S., Rastogi, A. and Juszczak, R., 2020. Review of top-of-canopy sun-induced fluorescence (SIF) studies from ground, UAV, airborne to spaceborne observations. *Sensors*, *20*(4), p.1144.

Berti Suman, A. and Alblas, E., 2023. Exploring Citizen Science over Time: Sensing, Technology and the Law. *Sustainability*, *15*(5), p.4496.

Bonney, R., Cooper, C.B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K.V. and Shirk, J., 2009. Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience*, *59*(11), pp.977-984.

Bonter, D.N. and Cooper, C.B., 2012. Data validation in citizen science: a case study from Project FeederWatch. *Frontiers in Ecology and the Environment*, *10*(6), pp.305-307.

Boote, J., Baird, W. and Beecroft, C., 2010. Public involvement at the design stage of primary health research: a narrative review of case examples. *Health policy*, *95*(1), pp.10-23.

Borre, J.V., Paelinckx, D., Mücher, C.A., Kooistra, L., Haest, B., De Blust, G. and Schmidt, A.M., 2011. Integrating remote sensing in Natura 2000 habitat monitoring: Prospects on the way forward. *Journal for Nature Conservation*, *19*(2), pp.116-125.

Bowman, N.D. and Keene, J.R., 2018. A layered framework for considering open science practices.

Boyle, T.J. and Sayer, J.A., 1995. Measuring, monitoring and conserving biodiversity in managed tropical forests. *The Commonwealth Forestry Review*, pp.20-25.

Butkevičienė, E., Skarlatidou, A., Balázs, B., Duží, B., Massetti, L., Tsampoulatidis, I. and Tauginienė, L., 2021. Citizen science case studies and their impacts on social innovation. *The science of citizen science*, p.309.

Camacho, A.E., 2020. De-and re-constructing public governance for biodiversity conservation. *Vand. L. Rev.*, *73*, p.1585.

Campbell, J.B. and Wynne, R.H., 2011. *Introduction to remote sensing*. Guilford Press.

Cappa, F., 2022. Big data from customers and non-customers through crowdsourcing, citizen science and crowdfunding. *Journal of Knowledge Management*, *26*(11), pp.308-323.

Ceballos, G., Ehrlich, P.R. and Dirzo, R., 2017. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the national academy of sciences*, *114*(30), pp.E6089-E6096.

Chandler, M., See, L., Copas, K., Bonde, A.M., López, B.C., Danielsen, F., Legind, J.K., Masinde, S., Miller-Rushing, A.J., Newman, G. and Rosemartin, A., 2017. Contribution of citizen science towards international biodiversity monitoring. *Biological Conservation*, *213*, pp.280-294.

Clements, J.C., 2017. Open access articles receive more citations in hybrid marine ecology journals. *Facets*, *2*(1), pp.1-14.

Collen, B., Ram, M., Zamin, T. and McRae, L., 2008. The tropical biodiversity data gap: addressing disparity in global monitoring. *Tropical Conservation Science*, *1*(2), pp.75-88.

Collin, A.M., Andel, M., James, D. and Claudet, J., 2019. The superspectral/hyperspatial worldview-3 as the link between spaceborne hyperspectral and airborne hyperspatial sensors: the case study of the complex tropical coast. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*.

Convention on Biological Diversity., 2011. Conference of the Parties Decision X/2: Strategic plan for biodiversity 2011–2020.

Convention on Biological Diversity., 2020. Aichi biodiversity targets. *[Online]*. Accessed 21st December 2020. Available at: https://www.cbd.int/sp/targets/

Cord, A.F., Brauman, K.A., Chaplin-Kramer, R., Huth, A., Ziv, G. and Seppelt, R., 2017. Priorities to advance monitoring of ecosystem services using earth observation. *Trends in ecology & evolution*, *32*(6), pp.416-428.

Cowie, H., Reis, S., Riddell, K. and Semple, S., 2014. Urban air quality citizen science. Project overview report. Accessed 2nd March 2021. Available at: http://nora.nerc.ac.uk/id/eprint/507912/1/N507912CR.pdf

Crall, A.W., Newman, G.J., Stohlgren, T.J., Holfelder, K.A., Graham, J. and Waller, D.M., 2011. Assessing citizen science data quality: an invasive species case study. *Conservation Letters*, *4*(6), pp.433-442.

Davies, L., Fradera, R., Riesch, H. and Lakeman-Fraser, P., 2016. Surveying the citizen science landscape: an exploration of the design, delivery and impact of citizen science through the lens of the Open Air Laboratories (OPAL) programme. *BMC ecology*, *16*(1), p.17.

de la Fuente, G. B., 2019. What is open science: Introduction. [Online]. Retrieved July 5, 2022, from https://www.fosteropenscience.eu/content/what-open-science-introduction

de Sherbinin, A., Bowser, A., Chuang, T.R., Cooper, C., Danielsen, F., Edmunds, R., Elias, P., Faustman, E., Hultquist, C., Mondardini, R. and Popescu, I., 2021. The critical importance of citizen science data. *Frontiers in Climate*, *3*, p.20.

Dickinson, J.L., Shirk, J., Bonter, D., Bonney, R., Crain, R.L., Martin, J., Phillips, T. and Purcell, K., 2012. The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the Environment*, *10*(6), pp.291-297.

Dickinson, J.L., Zuckerberg, B. and Bonter, D.N., 2010. Citizen science as an ecological research tool: challenges and benefits. *Annual review of ecology, evolution, and systematics*, *41*, pp.149-172.

Domroese, M.C. and Johnson, E.A., 2017. Why watch bees? Motivations of citizen science volunteers in the Great Pollinator Project. *Biological Conservation*, *208*, pp.40-47.

Druschke, C.G. and Seltzer, C.E., 2012. Failures of engagement: Lessons learned from a citizen science pilot study. *Applied Environmental Education & Communication*, *11*(3-4), pp.178-188.

Dubayah, R.O. and Drake, J.B., 2000. Lidar remote sensing for forestry. *Journal of Forestry*, *98*(6), pp.44-46.

Dusseux, P., Corpetti, T., Hubert-Moy, L. and Corgne, S., 2014. Combined use of multi-temporal optical and radar satellite images for grassland monitoring. *Remote Sensing*, *6*(7), pp.6163-6182.

Elliott, K.C. and Rosenberg, J., 2019. Philosophical foundations for citizen science. *Citizen Science: Theory and Practice*, *4*(1).

Ellis, S., 2003. Habitat quality and management for the northern brown argus butterfly *Aricia artaxerxes* (Lepidoptera: Lycaenidae) in North East England. *Biological Conservation*, *113*(2), pp.285-294.

Esri Inc., 2020. *ArcGIS Pro* (Version 2.5). Esri Inc. https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview.

Evans, D., 2012. Building the European union's Natura 2000 network. *Nature conservation*, *1*, pp.11-26.

Finch, T., ''Hanlon, N. and Dudley, S.P., 2017. Tweeting birds: online mentions predict future citations in ornithology. *Royal Society open science*, *4*(11), p.171371.

Fleishman, E. and Murphy, D.D., 2009. A realistic assessment of the indicator potential of butterflies and other charismatic taxonomic groups. *Conservation biology*, *23*(5), pp.1109-1116.

Follett, R. and Strezov, V., 2015. An analysis of citizen science based research: usage and publication patterns. *PloS one*, *10*(11), p.e0143687.

Fox, R., Dennis, E.B., Purdy, K.M., Middlebrook, I., Roy, D.B., Noble, D.G., Botham, M.S., and Bourn, N.A.D., 2023. The State of the UK's Butterflies 2022. Butterfly Conservation, Wareham, UK.

Fraisl, D., Campbell, J., See, L.,In, U., Wardlaw, J., Gold, M., Moorthy, I., Arias, R., Piera, J., Oliver, J.L. and Masó, J., 2020. Mapping citizen science contributions to the UN sustainable development goals. *Sustainability Science*, *15*(6), pp.1735-1751.

Fritz, S., Fonte, C.C. and See, L., 2017. The role of citizen science in earth observation. *Remote Sensing*, *9*(4), p.357.

Fuller, R.A., Lee, J.R. and Watson, J.E., 2014. Achieving open access to conservation science. *Conservation Biology*, *28*(6), pp.1550-1557.

Gaikwad, J. and Chavan, V., 2006. Open access and biodiversity conservation: challenges and potentials for the developing world. *Data Science Journal*, *5*, pp.1-17.

Gaines, W.L., 1999. *Monitoring biodiversity: quantification and interpretation* (Vol. 443). US Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Gallacher, J., 2007. The northern brown argus in23oil2323oil23 hi23oil23mannanshireshire. *Forth Naturalist and Historian*, *30*, p.73.

Geijzendorffer, I.R., Regan, E.C., Pereira, H.M., Brotons, L., Brummitt, N., Gavish, Y., Haase, P., Martin, C.S., Mihoub, J.B., Secades, C. and Schmeller, D.S., 2016. Bridging the gap between biodiversity data and policy reporting needs: An Essential Biodiversity Variables perspective. *Journal of Applied Ecology*, *53*(5), pp.1341-1350.

Geller, G.N., Halpin, P.N., Helmuth, B., Hestir, E.L., Skidmore, A., Abrams, M.J., Aguirre, N., Blair, M., Botha, E., Colloff, M. and Dawson, T., 2017. Remote sensing for biodiversity. In *The GEO handbook on biodiversity observation networks* (pp. 187-210). Springer, Cham.

Girardello, M., Martellos, S., Pardo, A. And Bertolino, S., 2018. Gaps in biodiversity occurrence information may hamper the achievement of international biodiversity targets: insights from a cross-taxon analysis. *Environmental Conservation*, *45*(4), pp.370-377.

Gong, K., 2022. Open science: The science paradigm of the new era. *Cultures of Science*, *5*(1), pp.3-9.

González-Oreja, J.A., Garbisu, C., Mijangos, I., Mendarte, S. and Albizu, I., 2013. Reducing costs in biodiversity monitoring: Shortcuts for plant diversity in meadows as a case study. *Ecological indicators*, *24*, pp.96-104.

Grainger, A., 2017. Citizen observatories and the new earth observation science. *Remote Sensing*, *9*(2), p.153.

Habel, J.C., Samways, M.J. and Schmitt, T., 2019. Mitigating the precipitous decline of terrestrial European insects: Requirements for a new strategy. *Biodiversity and Conservation*, *28*(6), pp.1343-1360.

Hagerman, S.M., Campbell, L.M., Gray, N.J. and Pelai, R., 2021. Knowledge production for target-based biodiversity governance. *Biological Conservation*, *255*, p.108980.

Hajjem, C., Harnad, S. and Gingras, Y., 2006. Ten-year cross-disciplinary comparison of the growth of open access and how it increases research citation impact. *arXiv preprint cs/0606079*.

Hallmann, C.A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., Stenmans, W., Müller, A., Sumser, H., Hörren, T. and Goulson, D., 2017. More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PloS one*, *12*(10), p.e0185809.

Hance, J., 2019. The great insect dying: Vanishing act in Europe and North America. *[Online]*. Accessed on 11th December 2020. Available at: https://news.mongabay.com/2019/06/the-great-insect-dying-vanishing-act-in-europe-and-north-america/

Hassan, A., Roberts, L. and Rodger, K., 2022. Corporate accountability for biodiversity and species extinction: Evidence from organisations reporting on their impacts on nature. *Business Strategy and the Environment*, *31*(1), pp.326-352.

Hayhow, D.B., Eaton, M.A., Stanbury, A.J., Burns ,F., Kirby, W.B., Bailey, N., Beckmann, B., Bedford, J., Boersch-Supan, P.H., Coomber, F., Dennis, E.B., Dolman, S.J., Dunn, E., Hall, J., Harrower, C., Hatfield, J.H., Hawley, J., Haysom, K., Hughes, J., Johns, D.G., Mathews, F., McQuatters-Gollop, A., Noble, D.G., Outhwaite, C.L., Pearce-Higgins, J.W., Pescott, O.L., Powney, G.D., and Symes, N., 2019. The State of Nature 2019. *The State of Nature partnership*. Accessed 4th January 2021. Available at: https://nbn.org.uk/wp-content/uploads/2019/09/State-of-Nature-2019-UK-full-report.pdf

Heberling, J.M., Miller, J.T., Noesgaard, D., Weingart, S.B. and Schigel, D., 2021. Data integration enables global biodiversity synthesis. *Proceedings of the National Academy of Sciences*, *118*(6), p.e2018093118.

Henle, K., Bauch, B., Auliya, M., Külvik, M', Pe'er, G., Schmeller, D.S. and Framstad, E., 2013. Priorities for biodiversity monitoring in Europe: A review of supranational policies and a novel scheme for integrative prioritization. *Ecological Indicators*, *33*, pp.5-18.

Henry, P.Y., Lengyel, S., Nowicki, P., Julliard, R., Clobert, J., Čelik, T., Gruber, B., Schmeller, D.S., Babij, V. and Henle, K., 2008. Integrating ongoing biodiversity monitoring: potential benefits and methods. *Biodiversity and conservation*, *17*, pp.3357-3382.

Hochkirch, A., Samways, M.J., Gerlach, J., Böhm, M., Williams, P., Cardoso, P., Cumberlidge, N., Stephenson, P.J., Seddon, M.B., Clausnitzer, V. and Borges, P.A., 2020. A strategy for the next decade to address data deficiency in neglected biodiversity. *Conservation Biology*.

Holbrook, J.B., 2019. Open science, open access, and the democratization of knowledge. *Issues in Science and Technology*, *35*(3), pp.26-28.

Hollow, B., Roetman, P.E., Walter, M. and Daniels, C.B., 2015. Citizen science for policy development: The case of koala management in South Australia. *Environmental Science & Policy*, *47*, pp.126-136.

Hooykaas, M.J., Schilthuizen, M., Aten, C., Hemelaar, E.M., Albers, C.J. and Smeets, I., 2019. Identification skills in biodiversity professionals and laypeople: A gap in species literacy. *Biological Conservation*, *238*, p.108202.

Hoye, T.T., Groom, Q., Juslén, A., Mandon, C., Dinesen, L., Rosenberg, A., Hendriks, R.J., Eggermont, H. and Vihervaara, P., 2022. Biodiversity monitoring knowledge gaps and research & innovation priorities.

Hunter, J., Alabri, A. and van Ingen, C., 2013. Assessing the quality and trustworthiness of citizen science data. *Concurrency and Computation: Practice and Experience*, *25*(4), pp.454-466.

Ichter, J., Evans, D. and Richard, D., 2014. Terrestrial habitat mapping in Europe: an overview. *Publications Office of the European Union, Luxembourg*.

Ignat, T. and Ayris, P., 2020. Built to last! Embedding open science principles and practice into European universities. *Insights*, *33*(1).

IPBES., 2019. Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. S. Díaz, J. Settele, E. S. Brondízio E.S., H. T. Ngo, M. Guèze, J. Agard, A. Arneth, P. Balvanera, K. A. Brauman, S. H. M. Butchart, K. M. A. Chan, L. A. Garibaldi, K. Ichii, J. Liu, S. M. Subramanian, G. F. Midgley, P. Miloslavich, Z. Molnár, D. Obura, A. Pfaff, S. Polasky, A. Purvis, J. Razzaque, B. Reyers, R. Roy Chowdhury, Y. J. Shin, I. J. Visseren-Hamakers, K. J. Willis, and C. N. Zayas (eds.). IPBES secretariat, Bonn, Germany. 56 pages.

JNCC., 2010. *Handbook for Phase 1 Habitat Survey – a Technique for Environmental Audit*. Joint Nature Conservation Committee, Peterborough

Johnston, A., Hochachka, W.M., Strimas-Mackey, M.E., Gutierrez, V.R., Robinson, O.J., Miller, E.T., Auer, T., Kelling, S.T. and Fink, D., 2020. Analytical guidelines to increase the value of citizen science data: using eBird data to estimate species occurrence. *bioRxiv*, p.574392.

Jones, B.L., Unsworth, R.K., McKenzie, L.J., Yoshida, R.L. and Cullen-Unsworth, L.C., 2018. Crowdsourcing conservation: The role of citizen science in securing a future for seagrass. *Marine pollution bulletin*, *134*, pp.210-215.

Khorram, S., van der Wiele, C.F., Koch, F.H., Nelson, S.A., Potts, M.D., Khorram, S., van der Wiele, C.F., Koch, F.H., Nelson, S.A. and Potts, M.D., 2016. Future trends in remote sensing. *Principles of applied remote sensing*, pp.277-285.

Kindsvater, H.K., Dulvy, N.K., Horswill, C., Juan-Jordá, M.J., Mangel, M. and Matthiopoulos, J., 2018. Overcoming the data crisis in biodiversity conservation. *Trends in ecology & evolution*, *33*(9), pp.676-688.

Klein, O., Hardwicke, T.E., Aust, F., Breuer, J., Danielsson, H., Hofelich Mohr, a., Ijzerman, H., Nilsonne, G., Vanpaemel, W. and Frank, M.C., 2018. A practical guide for transparency in psychological science. *Collabra: Psychology*, *4*(1), pp.1-15.

Klemas, V.V., 2001. Remote sensing of landscape-level coastal environmental indicators. *Environmental management*, *27*, pp.47-57.

Kosmala, M., Wiggins, A., Swanson, A. and Simmons, B., 2016. Assessing data quality in citizen science. *Frontiers in Ecology and the Environment*, *14*(10), pp.551-560.

Kühl, H.S., Bowler, D.E., Bösch, L., Bruelheide, H., Dauber, J., Eichenberg, D., Eisenhauer, N., Fernández, N., Guerra, C.A., Henle, K. and Herbinger, I., 2020. Effective biodiversity monitoring needs a culture of integration. *One Earth*, *3*(4), pp.462-474.

Kullenberg, C. And Kasperowski, D., 2016. What is citizen science?–A scientometric meta-analysis. *PloS one*, *11*(1), p.e0147152.

Lamb, C.T., Gilbert, S.L. and Ford, A.T., 2018. Tweet success? Scientific communication correlates with increased citations in Ecology and Conservation. *PeerJ*, *6*, p.e4564.

Lamb, E.G., Bayne, E., Holloway, G., Schieck, J., Boutin, S., Herbers, J. and Haughland, D.L., 2009. Indices for monitoring biodiversity change: Are some more effective than others?. *Ecological indicators*, *9*(3), pp.432-444.

Lambers, K., Verschoof-van der Vaart, W.B. and Bourgeois, Q.P., 2019. Integrating remote sensing, machine learning, and citizen science in Dutch archaeological prospection. *Remote Sensing*, *11*(7), p.794.

Land-Zandstra, A., Agnello, G. and Gültekin, Y.S., 2021. Participants in citizen science. *The science of citizen science*, *243*, p.259.

Lee, W., McGlone, M. and Wright, E., 2005. Biodiversity inventory and monitoring: a review of national and international systems and a proposed framework for future biodiversity monitoring by the Department of Conservation. *Landcare Research contract report LC0405/122*.

Levin, N., Leonelli, S., Weckowska, D., Castle, D. and Dupré, J., 2016. How do scientists define openness? Exploring the relationship between open science policies and research practice. *Bulletin of science, technology & society*, *36*(2), pp.128-141.

Lewis, P. and Hancock, S., 2007. LiDAR for vegetation applications. *UCL, Gower St, London, UK*. Available at: http://www2.geog.ucl.ac.uk/~mdisney/teaching/PPRS/PPRS_6/LiDARforVegetationApplications.pdf

Lindenmayer, D.B., Gibbons, P., Bourke, M.A.X., Burgman, M., Dickman, C.R., Ferrier, S., Fitzsimons, J., Freudenberger, D., Garnett, S.T., Groves, C. and Hobbs, R.J., 2012. Improving biodiversity monitoring. *Austral Ecology*, *37*(3), pp.285-294.

Luque, S., Pettorelli, N., Vihervaara, P. and Wegmann, M., 2018. Improving biodiversity monitoring using satellite remote sensing to provide solutions towards the 2020 conservation targets. *Methods in Ecology and Evolution*, *9*(8), pp.1784-1786.

MacCallum, C.J., 2007. When is open access not open access?. *25oil 25oil*, *5*(10), p.e285.

Maes, J., Egoh, B., Willemen, L., Liquete, C., Vihervaara, P., Schägner, J.P., Grizzetti, B., Drakou, E.G., La Notte, A., Zulian, G. and Bouraoui, F., 2012. Mapping ecosystem services for policy support and decision making in the European Union. *Ecosystem services*, *1*(1), pp.31-39.

Magnificent Meadows., n.d. Common Rock-rose *Helianthemum nummularium*. *[Online]*. Accessed 17th February 2021. Available at: http://www.magnificentmeadows.org.uk/assets/pdfs/Common_rockrose.pdf

Magurran, A.E., Baillie, S.R., Buckland, S.T., Dick, J.M., Elston, D.A., Scott, E.M., Smith, R.I., Somerfield, P.J. and Watt, A.D., 2010. Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. *Trends in ecology & evolution*, *25*(10), pp.574-582.

Malhi, Y., 2017. The concept of the Anthropocene. *Annual Review of Environment and Resources*, *42*, pp.77-104.

Mallet, J., Wynne, I.R. and Thomas, C.D., 2011. Hybridisation and climate change: brown argus butterflies in Britain (Polyommatus subgenus *Aricia*). *Insect Conservation and Diversity*, *4*(3), pp.192-199.

Marshall, E., Wintle, B.A., Southwell, D. and Kujala, H., 2020. What are we measuring? A review of metrics used to describe biodiversity in offsets exchanges. *Biological Conservation*, *241*, p.108250.

Martin, L.J., Blossey, B. and Ellis, E., 2012. Mapping where ecologists work: biases in the global distribution of terrestrial ecological observations. *Frontiers in Ecology and the Environment*, *10*(4), pp.195-201.

Martin, V.Y., 2017. Citizen science as a means for increasing public engagement in science: Presumption or possibility?. *Science Communication*, *39*(2), pp.142-168.

Mazumdar, S., Wrigley, S. and Ciravegna, F., 2017. Citizen science and crowdsourcing for earth observations: An analysis of stakeholder opinions on the present and future. *Remote Sensing*, *9*(1), p.87.

McKinley, D.C., Miller-Rushing, A.J., Ballard, H.L., Bonney, R., Brown, H., Evans, D.M., French, R.A., Parrish, J.K., Phillips, T.B., Ryan, S.F. and Shanley, L.A., 2015. Investing in citizen science can improve natural resource management and environmental protection. *Issues in Ecology*, *2015*(19), pp.1-27.

Meneghini, A., 2019. An Evaluation of Sentinel-1 and Sentinel-2 for Land Cover Classification.

Mihoub, J.B., Henle, K., Titeux, N., Brotons, L., Brummitt, N.A. and Schmeller, D.S., 2017. Setting temporal baselines for biodiversity: the limits of available monitoring data for capturing the full impact of anthropogenic pressures. *Scientific Reports*, *7*, p.41591.

Miller-Rushing, A.J., Primack, R.B., Bonney, R.I.C.K. and Albee, E.M.M.A., 2020. The history of citizen science in Ecology and Conservation. *Lepczyk, CA*, pp.17-24.

Montes, E., Lefcheck, J.S., Guerra-Castro, E., Klein, E., Kavanaugh, M.T., de Azevedo Mazzuco, A.C., Bigatti, G., Cordeiro, C.A., Simoes, N., Macaya, E.C. and Moity, N., 2021. OPTIMIZING LARGE-SCALE BIODIVERSITY SAMPLING EFFORT. *Oceanography*, *34*(2), pp.80-91.

Muhamedieva, D.T. and Niyozmatova, N.A., 2023. Monitoring of biodiversity of water communities. In *E3S Web of Conferences* (Vol. 411, p. 02042). EDP Sciences.

Nagendra, H., Lucas, R., Honrado, J.P., Jongman, R.H., Tarantino, C., Adamo, M. and Mairota, P., 2013. Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats. *Ecological Indicators*, *33*, pp.45-59.

Natural England., 2020. Climate Change Adaptation Manual (NE751). *[Online]*. Accessed 11th December 2020. Available at: http://publications.naturalengland.org.uk/publication/5679197848862720

NatureScot., 2011. Guide to species-rich grasslands. *[Online]*. Accessed 7th December 2020. Available at: https://www.nature.scot/guide-types-species-rich-grassland

Navarro, L.M., Fernández, N., Guerra, C., Guralnick, R., Kissling, W.D., Londoño, M.C., Muller-Karger, F., Turak, E., Balvanera, P., Costello, M.J. and Delavaud, A., 2017. Monitoring biodiversity change through effective global coordination. *Current opinion in environmental sustainability*, *29*, pp.158-169.

Navarro-Perez, M. and Tidball, K.G., 2012. Challenges of Biodiversity Education: A Review of Education Strategies for Biodiversity Education. *International Electronic Journal of Environmental Education*, *2*(1), pp.13-30.

Newman, G., Wiggins, A., Crall, A., Graham, E., Newman, S. and Crowston, K., 2012. The future of citizen science: emerging technologies and shifting paradigms. *Frontiers in Ecology and the Environment*, *10*(6), pp.298-304.

Neylon, C. and Wu, S., 2009. Open Science: tools, approaches, and implications. In *Biocomputing 2009* (pp. 540-544).

Nicholas, D., Boukacem-Zeghmouri, C., Abrizah, A., Rodríguez-Bravo, B., Xu, J., Świgon, M., Watkinson, A. and Herman, E., 2019. Open science from the standpoint of the new wave of researchers: Views from the scholarly frontline. *Information Services & Use*, (Preprint), pp.1-6.

Nichols, J.D. and Williams, B.K., 2006. Monitoring for conservation. *Trends in ecology & evolution*, *21*(12), pp.668-673.

Nosek, B.A., Beck, E.D., Campbell, L., Flake, J.K., Hardwicke, T.E., Mellor, D.T., van't Veer, A.E. and Vazire, S., 2019. Preregistration is hard, and worthwhile. *Trends in cognitive sciences*, *23*(10), pp.815-818.

Nugent, J., 2018. iNaturalist: citizen science for 21st-century naturalists. *Science Scope*, *41*(7), p.12.

Obanawa, H., Yoshitoshi, R., Watanabe, N. and Sakanoue, S., 2020. Portable LiDAR-Based Method for Improvement of Grass Height Measurement Accuracy: Comparison with SfM Methods. *Sensors*, *20*(17), p.48'9.

O'Connor, B., Secades, C., Penner, J., Sonnenschein, R., Skidmore, A., Burgess, N.D. and Hutton, J.M., 2015. Earth observation as a tool for tracking progress towards the Aichi Biodiversity Targets. *Remote sensing in ecology and conservation*, *1*(1), pp.19-'8.

Oliver, R.Y., Meyer, C., Ranipeta, A., Winner, K. and Jetz, W., 2021. Global and national trends, gaps, and opportunities in documenting and monitoring species distributions. *PloS Biology*, *19*(8), p.e3001336.

O'Mara, F.P., 2012. The role of grasslands in food security and climate change. *Annals of botany*, *110*(6), pp.1263-1270.

Ondei, S., Brook, B.W. and Buettel, J.C., 2018. Nature's untold stories: an overview on the availability and type of on-line data on long-term biodiversity monitoring. *Biodiversity and Conservation*, *27*(11), pp.2971-2987.

Oostermeijer, J.G.B. and Van Swaay, C.A.M., 1998. The relationship between butterflies and environmental indicator values: a tool for conservation in a changing landscape. *Biological conservation*, *86*(3), pp.271-280.

Ota, T., Ogawa, M., Shimizu, K., Kajisa, T., Mizoue, N., Yoshida, S., Takao, G., Hirata, Y., Furuya, N., Sano, T. and Sokh, H., 2015. Aboveground biomass estimation using structure from motion approach with aerial photographs in a seasonal tropical forest. *Forests*, *6*(11), pp.3882-3898.

Perino, A., Pereira, H.M., Felipe-Lucia, M., Kim, H., Kühl, H.S., Marselle, M.R., Meya, J.N., Meyer, C., Navarro, L.M., van Klink, R. and Albert, G., 2022. Biodiversity post-2020: Closing the gap between global targets and national-level implementation. *Conservation letters*, *15*(2), p.e12848.

Petrou, Z.I., Manakos, I. and Stathaki, T., 2015. Remote sensing for biodiversity monitoring: a review of methods for biodiversity indicator extraction and assessment of progress towards international targets. *Biodiversity and conservation*, *24*, pp.2333-2363.

Pettorelli, N., Safi, K. and Turner, W., 2014. Satellite remote sensing, biodiversity research and conservation of the future.

Plantlife., 2018. Grasslands action plan. *[Online]*. Accessed 7th December 2020. Available at: https://www.plantlife.org.uk/application/files/2315/3087/2058/Grasslands_action_plan_-_Plantlife.pdf

Pocock, M.J., Newson, S.E., Henderson, I.G., Peyton, J., Sutherland, W.J., Noble, D.G., Ball, S.G., Beckmann, B.C., Biggs, J., Brereton, T. and Bullock, D.J., 2015. Developing and enhancing biodiversity monitoring programmes: a collaborative assessment of priorities. *Journal of Applied Ecology*, *52*(3), pp.686-695.

Pocock, M.J., Roy, H.E., August, T., Kuria, A., Barasa, F., Bett, J., Githiru, M., Kairo, J., Kimani, J., Kinuthia, W. and Kissui, B., 2019. Developing the global potential of citizen science: Assessing opportunities that benefit people, society and the environment in East Africa. *Journal of Applied Ecology*, *56*(2), pp.274-281.

Quaye-Ballard, N.L., Asenso-Gyambibi, D. and Quaye-Ballard, J., 2020. Unmanned Aerial Vehicle for Topographical Mapping of Inaccessible Land Areas in Ghana: A Cost-Effective Approach.

Raddick, M.J., Prather, E.E. and Wallace, C.S., 2019. Galaxy zoo: Science content knowledge of citizen scientists. *Public Understanding of Science*, *28*(6), pp.636-651.

Rapinel, S., Mony, C., Lecoq, L., Clement, B., Thomas, A. and Hubert-Moy, L., 2019. Evaluation of Sentinel-2 time-series for mapping floodplain grassland plant communities. *Remote sensing of environment*, *223*, pp.115-129.

Reddy, C.S., 2021. Remote sensing of biodiversity: what to measure and monitor from space to species?. *Biodiversity and Conservation*, *30*(10), pp.2617-2631.

Reddy, C.S., Kurian, A., Srivastava, G., Singhal, J., Varghese, A.O., Padalia, H., Ayyappan, N., Rajashekar, G., Jha, C.S. and Rao, P.V.N., 2021. Remote sensing enabled essential biodiversity variables for biodiversity assessment and monitoring: technological advancement and potentials. *Biodiversity and Conservation*, *30*, pp.1-14.

Reinermann, S., Asam, S. and Kuenzer, C., 2020. Remote Sensing of Grassland Production and Management—A Review. *Remote Sensing*, *12*(12), p.1949.

Roche, D.G., O'Dea, R.E., Kerr, K.A., Rytwinski, T., Schuster, R., Nguyen, V.M., Young, N., Bennett, J.R. and Cooke, S.J., 2022. Closing the knowledge-action gap in conservation with open science. *Conservation Biology*, *36*(3), p.e13835.

Rodwell, J.S. ed., 1998. *British Plant Communities: Volume 3, Grasslands and Montane Communities* (Vol. 3). Cambridge University Press.

Rotman, D., Preece, J., Hammock, J., Procita, K., Hansen, D., Parr, C., Lewis, D. and Jacobs, D., 2012, February. Dynamic changes in motivation in collaborative citizen-science projects. In *Proceedings of the ACM 2012 conference on computer supported cooperative work* (pp. 217-226).

Roy, P.S., Behera, M.D. and Srivastav, S.K., 2017. Satellite remote sensing: sensors, applications and techniques. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, *87*, pp.465-472.

Ruckelshaus, M.H., Jackson, S.T., Mooney, H.A., Jacobs, K.L., Kassam, K.A.S., Arroyo, M.T., Báldi, A., Bartuska, A.M., Boyd, J., Joppa, L.N. and Kovács-Hostyánszki, A., 2020. The IPBES global assessment: pathways to action. *Trends in Ecology & Evolution*, *35*(5), pp.407-414.

Rülke, J., Rieckmann, M., Nzau, J.M. and Teucher, M., 2020. How Ecocentrism and Anthropocentrism Influence Human–Environment Relationships in a Kenyan Biodiversity Hotspot. *Sustainability*, *12*(19), p.8213.

Rydén, O., Zizka, A., Jagers, S.C., Lindberg, S.I. and Antonelli, A., 2020. Linking democracy and biodiversity conservation: Empirical evidence and research gaps. *Ambio*, *49*(2), pp.419-433.

Saluja, M.K. and Thakur, S.S., 2020. Open Source Software Based Education and Training Framework for Software Engineering Education. *Solid State Technology*, *63*(5), pp.9633-9645.

Sánchez-Bayo, F. and Wyckhuys, K.A., 2019. Worldwide decline of the entomofauna: A review of its drivers. *Biological conservation*, *232*, pp.8-27.

Sarafoglou, A., Kovacs, M., Bakos, B., Wagenmakers, E.J. and Aczel, B., 2022. A survey on how preregistration affects the research workflow: Better science but more work. *Royal Society Open Science*, *9*(7), p.211997.

Schröter, M., Kraemer, R., Mantel, M., Kabisch, N., Hecker, S., Richter, A., Neumeier, V. And Bonn, A., 2017. Citizen science for assessing ecosystem services: status, challenges and opportunities. *Ecosystem Services*, *28*, pp.80-94.

Schwieder, M., Buddeberg, M., Kowalski, K., Pfoch, K., Bartsch, J., Bach, H., Pickert, J. and Hostert, P., 2020. Estimating Grassland Parameters from Sentinel-2: A Model Comparison Study. *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, *88*(5), pp.379-390.

Scudder, G.G., 2017. The importance of insects. *Insect biodiversity: science and society*, *1*, pp.9-43.

Seddon, N., Chausson, A., Berry, P., Girardin, C.A., Smith, A. and Turner, B., 2020. Understanding the value and limits of nature-based solutions to climate change and other global challenges. *Philosophical Transactions of the Royal Society B*, *375*(1794), p.20190120.

See, L., Bayas, J.C.L., Lesiv, M., Schepaschenko, D., Danylo, O., McCallum, I., Dürauer, M., Georgieva, I., Domian, D., Fraisl, D. and Hager, G., 2022. Lessons learned in developing reference data sets with the contribution of citizens: the Geo-Wiki experience. *Environmental Research Letters*, *17*(6), p.065003.

Seibold, S., Gossner, M.M., Simons, N.K., Blüthgen, N., Müller, J., Ambarlı, D., Ammer, C., Bauhus, J., Fischer, M., Habel, J.C. and Linsenmair, K.E., 2019. Arthropod decline in grasslands and forests is associated with landscape-level drivers. *Nature*, *574*(7780), pp.671-674.

Shellswell, C.H., Chant, J.J., Jefferson, R.G., Le Bas, B., Edwards, J. and Parton, C., 2016. Restoration of existing lowland grassland – timescales to achieve favourable condition. Plantlife, Salisbury. Accessed 25th January 2021. Available at: http://www.magnificentmeadows.org.uk/assets/pdfs/Lowland_grassland_timescales_for_recovery_advisory_note_FINAL-Design.pdf

Siddig, A.A., 2019. Why is biodiversity data-deficiency an ongoing conservation dilemma in Africa?. *Journal for Nature Conservation*, *50*, p.125719.

Smith, Elta, Sarah Parks, Salil Gunashekar, Catherine A. Lichten, Anna Knack, and Catriona Manville, Open Science: The citizen's role and contribution to research. Santa Monica, CA: RAND Corporation, 2017. https://www.rand.org/pubs/perspectives/PE246.html.

Smith, T., Paavola, J. and Holmes, G., 2019. Corporate reporting and conservation realities: Understanding differences in what businesses say and do regarding biodiversity. *Environmental Policy and Governance*, *29*(1), pp.3-13.

Sommervold, O., Gazzea, M. and Arghandeh, R., 2023. A survey on sar and optical satellite image registration. *Remote Sensing*, *15*(3), p.850.

Stephenson, P.J., 2020. Technological advances in biodiversity monitoring: applicability, opportunities and challenges. *Current Opinion in Environmental Sustainability*, *45*, pp.36-41.

Stephenson, P.J. and Stengel, C., 2020. An inventory of biodiversity data sources for conservation monitoring. *PLoS One*, *15*(12), p.e0242923.

Stephenson, P.J., Brooks, T., Butchart, S., Fegraus, E., Geller, G., Hoft, R., Hutton, J., Kingston, N., Long, B. and McRae, L., 2017. Priorities for big biodiversity data. *Frontiers in Ecology and the Environment*, *15*(3), pp.124-125.

Sterling, E.J., Betley, E., Sigouin, A., Gomez, A., Toomey, A., Cullman, G., Malone, C., Pekor, A., Arengo, F., Blair, M. and Filardi, C., 2017. Assessing the evidence for stakeholder engagement in biodiversity conservation. *Biological conservation*, *209*, pp.159-171.

Stork, N.E., 2018. How many species of insects and other terrestrial arthropods are there on Earth?. *Annual review of entomology*, *63*, pp.31-45.

Sugai, L.S., 2020. Pandemics and the Need for Automated Systems for Biodiversity Monitoring. *The Journal of Wildlife Management*.

Suter, S., Barrett, B. and Welden, N., 2023. Do biodiversity monitoring citizen science surveys meet the core principles of open science practices?. *Environmental Monitoring and Assessment*, *195*(2), pp.1-14.

Swanson, A., Kosmala, M., Lintott, C. and Packer, C., 2016. A generalized approach for producing, quantifying, and validating citizen science data from wildlife images. *Conservation Biology*, *30*(3), pp.520-531.

Taylor, B., Chapron, G., Kopnina, H., Orlikowska, E., Gray, J. and Piccolo, J.J., 2020*a*. The need for ecocentrism in biodiversity conservation. *Conservation Biology*, *34*(5), pp.1089-1096.

Taylor, B., Wright, J. and LeVasseur, T., 2020*b*. Dark green humility: religious, psychological., and affective attributes of proenvironmental behaviors. *Journal of Environmental Studies and Sciences*, *10*(1), pp.41-56.

Tengö, M., Austin, B.J., Danielsen, F. and Fernández-Llamazares, Á., 2021. Creating synergies between citizen science and Indigenous and local knowledge. *BioScience*, *71*(5), pp.503-518.

The Wildlife Trusts., n.d. Common Rock-rose. *[Online].* Accessed 17th February 2021. Available at: https://www.wildlifetrusts.org/wildlife-explorer/wildflowers/common-rock-rose

Theobald, E.J., Ettinger, A.K., Burgess, H.K., DeBey, L.B., Schmidt, N.R., Froehlich, H.E., Wagner, C., HilleRisLambers, J., Tewksbury, J., Harsch, M.A. and Parrish, J.K., 2015. Global change and local solutions: Tapping the unrealized potential of citizen science for biodiversity research. *Biological Conservation*, *181*, pp.236-244.

Thomas, C.D., Franco, A.M. and Hill, J.K., 2006. Range retractions and extinction in the face of climate warming. *Trends in Ecology & Evolution*, *21*(8), pp.415-416.

Thornley, R., Gerard, F.F., White, K. and Verhoef, A., 2022. Intra-annual taxonomic and phenological drivers of spectral variance in grasslands. *Remote Sensing of Environment*, *271*, p.112908.

Toomey, A.H. and Domroese, M.C., 2013. Can citizen science lead to positive conservation attitudes and behaviors?. *Human Ecology Review*, pp.50-62.

Troudet, J., Grandcolas, P., Blin, A., Vignes-Lebbe, R. and Legendre, F., 2017. Taxonomic bias in biodiversity data and societal preferences. *Scientific reports*, *7*(1), pp.1-14.

UK Biodiversity Action Plan. 2008. Priority Habitat Descriptions. BRIG (ed. Ant Maddock).

UKBMS., 2019 G1: monitoring butterfly numbers by the transect method - summary information for recorders. *[Online].* Accessed 4th March 2021. Available at: https://www.ukbms.org/Downloads/UKBMS%20G1%20Transect%20monitoring%20summary.pdf

Valdez, J.W., Callaghan, C.T., Junker, J., Purvis, A., Hill, S.L. and Pereira, H.M., 2023. The undetectability of global biodiversity trends using local species richness. *Ecography*, *2023*(3), p.e06604.

Vali, A., Comai, S. and Matteucci, M., 2020. Deep learning for land use and land cover classification based on hyperspectral and multispectral earth observation data: A review. *Remote Sensing*, *12*(15), p.2495.

Van Noorden, R., 2014. Confusion over open-data rules. *Nature*, *515*(7528), p.478.

Van Swaay, C.A., Dennis, E.B., Schmucki, R., Sevilleja, C.G., Balalaikins, M., Botham, M., Bourn, N., Brereton, T., Cancela, J.P., Carlisle, B. and Chambers, P., 2019. The EU Butterfly Indicator for Grassland species: 1990-2017: Technical Report. *Butterfly Conservation Europe & ABLE/eBMS. www. butterfly-monitoring. net*.Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., Samson, R. and Wagenknecht, K., 2021. The Science of Citizen Science Evolves. *Chapter 1. in Vohland, K. et al (Eds.).(2021) The Science of Citizen Science. Springer. https://doi. org/10.1007/978-3-030-58278-4. pp 1-12.*

Viqueira, J.R., Villarroya, S., Mera, D. and Taboada, J.A., 2020. Smart environmental data infrastructures: Bridging the gap between earth sciences and citizens. *Applied Sciences*, *10*(3), p.856.

Wachendorf, M., Fricke, T. and Möckel, T., 2018. Remote sensing as a tool to assess botanical composition, structure, quantity and quality of temperate grasslands. *Grass and Forage Science*, *73*(1), pp.1-14.

Walker, K.J., Stevens, P.A., Stevens, D.P., Mountford, J.O., Manchester, S.J. and Pywell, R.F., 2004. The restoration and re-creation of species-rich lowland grassland on land formerly managed for intensive agriculture in the UK. *Biological conservation*, *119*(1), pp.1-18.

Wallington, T.J., Hobbs, R.J. and Moore, S.A., 2005. Implications of current ecological thinking for biodiversity conservation: a review of the salient issues. *Ecology and Society*, *10*(1).

Walton, P., Eaton, M., Stanbury, A., Hayhow, D., Brand, A., Brooks, S., Collin, S., Duncan, C., Dundas, C., Foster, S., Hawley, J., Kinninmonth, A., Leatham, S., Nagy-Vizitiu, A., Whyte, A., Williams, S. and Wormald K., 2019. The State of Nature Scotland 2019. *The State of Nature partnership*. Accessed 4th January 2021. Available at: https://www.nature.scot/sites/default/files/2019-10/State-of-nature-Report-2019-Scotland-full-report.pdf

Warner, K.A., Lowell, B., Timme, W., Shaftel, E. and Hanner, R.H., 2019. Seafood sleuthing: How citizen science contributed to the largest market study of seafood mislabeling in the US and informed policy. *Marine Policy*, *99*, pp.304-311.

Wessels, P., Moran, N., Johnston, A. and Wang, W., 2019. Hybrid expert ensembles for identifying unreliable data in citizen science. *Engineering Applications of Artificial Intelligence*, *81*, pp.200-212.

Wetzel, F.T., Bingham, H.C., Groom, Q., Haase, P., Kõljalg, U., Kuhlmann, M., Martin, C.S., Penev, L., Robertson, T., Saarenmaa, H. and Schmeller, D.S., 2018. Unlocking biodiversity data: Prioritization and filling the gaps in biodiversity observation data in Europe. *Biological conservation*, *221*, pp.78-85.

Wiggins, A. and Crowston, K., 2012, January. Goals and tasks: Two typologies of citizen science projects. In *2012 45th Hawaii International Conference on System Sciences* (pp. 3426-3435). IEEE.

Williams, M., Bagwell, J. and Zozus, M.N., 2017. Data management plans: the missing perspective. *Journal of biomedical informatics*, *71*, pp.130-142.

Wood, E.M., Pidgeon, A.M., Radeloff, V.C. and Keuler, N.S., 2012. Image texture as a remotely sensed measure of vegetation structure. *Remote Sensing of Environment*, *121*, pp.516-526.

Worthington, J.P., Silvertown, J., Cook, L., Cameron, R., Dodd, M., Greenwood, R.M., McConway, K. and Skelton, P., 2012. Evolution MegaLab: a case study in citizen science methods. *Methods in Ecology and Evolution*, *3*(2), pp.303-309.

Yoccoz, N.G., Nichols, J.D. and Boulinier, T., 2001. Monitoring of biological diversity in space and time. *Trends in ecology & evolution*, *16*(8), pp.446-453.

Zlinszky, A., Heilmeier, H., Balzter, H., Czúcz, B. and Pfeifer, N., 2015. Remote sensing and GIS for habitat quality monitoring: New approaches and future research.

Zlinszky, A., Schroiff, A., Kania, A., Deák, B., Mücke, W., Vári, Á., Székely, B. and Pfeifer, N., 2014. Categorizing grassland vegetation with full-waveform airborne laser scanning: A feasibility study for detecting Natura 2000 habitat types. *Remote Sensing*, *6*(9), pp.8056-8087.

Chapter 2. Specific Targets for Biodiversity Monitoring

2.1 A Case-Study for Addressing the Thesis Aims

From the overarching aims of the thesis, I intended to create an OS biodiversity monitoring tool that combined CS and RS. This thesis will focus on biodiversity in the UK, not only targeting overlooked biodiversity components but concentrate efforts in one of the world's most nature-depleted countries, where previous anthropogenic activities have devastated the natural environment. A specific biodiversity case-study was focused on for this aim that was both identified from the literature and discussed with Stakeholder and UK NGO, Butterfly Conservation. From assessment of the literature, grasslands are consistently mentioned as a habitat that has been overlooked in global monitoring. This could also explain the lack of attention invertebrates have received as well, due to their synergistic relationship. It is pertinent to address the lack of monitoring of invertebrates and grassland habitats, due to their association and lack of attention.

Through the creation of the combined CS RS tool, and associated research, OS will be utilised where possible, and assessed in its success, thus reaching the first thesis aim. Specifically, the tool will be applied in mapping SRGs across the UK and to identify possible locations for the vulnerable Northern Brown Argus Butterfly, achieving the second thesis aim. As Butterfly Conservation, and other UK NGOs and departments, such as Plantlife and NatureScot, have specific targets for this case-study, ensuring stakeholder engagement and open science methodologies throughout the thesis will increase the capacity for public interactions and knowledge exchange.

2.2 Species-Rich Grasslands in the UK

Agricultural intensification in the UK has led to large declines in many natural and semi-natural habitats across the country (Walker *et al.*, 2004). One of the most affected habitats is species-rich grasslands (SRGs), which now make up less than 1% of the UK's land cover having been reduced by more than 97% since the last century (Hayhow *et al.*, 2019; Plantlife, 2018). SRGs are defined as a semi-natural landscape of native ancient communities, created with few agricultural inputs and, as such, result in a high array of biodiversity (usually >12 species per one square metre) (NatureScot, 2011; Rodwell, 1998). These grasslands may often be described as unimproved or, on occasion, semi-improved in relation to their agricultural enhancement (JNCC, 2010). The UK's remaining SRGs are still threatened by continued agricultural improvement, over- or under-grazing, pollution, afforestation, development, neglect, and climate change (UK Biodiversity Action Plan, 2008). Conserving SRGs is, therefore, critical to ensure no further losses are experienced.

Grasslands are crucial carbon sinks that help mitigate factors that influence global warming and, as such, provide essential ecosystem services to society (O'Mara, 2012). The loss of these habitats entirely would also result in irreversible damage to the UK's already vulnerable wildlife (Walton *et al.*, 2019). The diversity of flora and fauna supported by SRGs emphasises the vital importance of these habitats. Species heavily associated with SRGs are those in the Insecta class. Many insect habitats, including SRGs, are rapidly declining and, subsequently, insect population decreases have followed (Habel *et al.,* 2019). For example, land use changes in grasslands in Germany have resulted in a 67% decline in insect biomass and a 78% decline in abundance over a 9-year period (Seibold *et al.,* 2019). Declines are also prominent in protected areas (75% in less than 30 years in Germany), where conservation of related habitats should limit this (Hallmann *et al.,* 2017). These changes are not restricted spatially and have led to an estimate of 40% of insect species being at risk of extinction across the world (Sánchez-Bayo and Wyckhuys, 2019).

These insect declines have gained large attention in recent years after being long overlooked, due to their seemingly high abundance and rapid breeding. Potentially, lack of interest in these often-misunderstood species could have led to their disregard in monitoring, with only 0.8% of described species being assessed by the IUCN (Hance, 2019). However, the importance of insect species is now pronounced due to our increased understanding of their imperative roles in all life processes. These roles range from prey species to their pollination services which humans are heavily reliant on (to name only a few) (Scudder, 2017). Evidence suggests that insect declines are largest in grassland habitats and, therefore, targeting SRGs in conservation programmes has the potential to support a large array of species, many of which need protecting too (Sánchez-Bayo and Wyckhuys, 2019).

Assessment of the impact of habitat loss on insect orders has revealed that the Lepidoptera has been found to be one of the worst affected, for example, across the UK more than 65% of butterfly species' populations have declined over 43 years of surveying (Hance, 2019). These population decreases are associated with agricultural intensification and the fragmentation of their (largely grassland) habitats (Habel *et al.,* 2019). Butterfly populations are also essential indicator species, noted for their sensitivity to environmental or habitat changes, such as eutrophication, water depletion, acidification of soils, and global warming (Fleishman and Murphy, 2009; Oostermeijer and Van Swaay, 1998). Changes in these populations can, therefore, be indicative of wider issues, including deteriorating habitat condition and predicting other species declines. For example, 'The EU Butterfly Indicator for Grassland Species' was devised using a suite of butterflies associated with grasslands, that may be used to determine the wider state of biodiversity linked to these habitats. Furthermore, butterflies are one of the largest taxa of insects that have established monitoring schemes, making monitoring of wider biodiversity already more feasible (Van Swaay *et al.,* 2019). Identifying species that have this indicator potential of priority habitats, such as SRGs, would be hugely beneficial.

2.2.1 Priority Species of Species-Rich Grasslands

A species of conservation importance that is commonly associated with species-rich grasslands is the Northern Brown Argus butterfly, *Aricia artaxerxes*. In the UK, *A. artaxerxes* has seen a rapid reduction in population size and distribution since the 1970s (-57% and -56% respectively) (Fox *et al.,* 2023). *A. artaxerxes* is now heavily restricted to the Northeast of England and East of Scotland, with some occurrences in Scotland's Southwest, due largely to climate change shifting the species' range and falling habitat quality (Natural England, 2020). These worrying statistics have led the *A. artaxerxes* to be classed as UK Biodiversity Action Plan (BAP) priority species and of high conservation importance for the UK non-governmental organisation (NGO), Butterfly Conservation (Gallacher, 2007). If the species is at threat from increased climate change further restricting its range, ensuring its habitat is in favourable condition is fundamental through enhanced monitoring and management.

A. artaxerxes' distribution is greatly dependent on its larval food plant species *Helianthemum nummularium* (Common Rock-rose) which is found mostly on calcareous species-rich grasslands (although is able to grow on some acid and neutral grasslands too) (Figure 2-1). *H. nummularium* is currently not listed as vulnerable across the UK, especially in Scotland where it is listed as common. However, as calcareous grasslands have declined by 80%, further losses to these habitats could pose a risk to *H. nummularium* as well (The Wildlife Trusts, n.d.; Magnificent Meadows, n.d.). Requirements of calcareous SRGs for *A. artaxerxes,* besides the presence of *H. nummularium,* also include southerly facing slopes up to 350 m high in altitude and sward heights of ideally more than 10 cm, but at least 6 cm, indicating a light grazing regime is most beneficial for management of this habitat (Ellis, 2003).

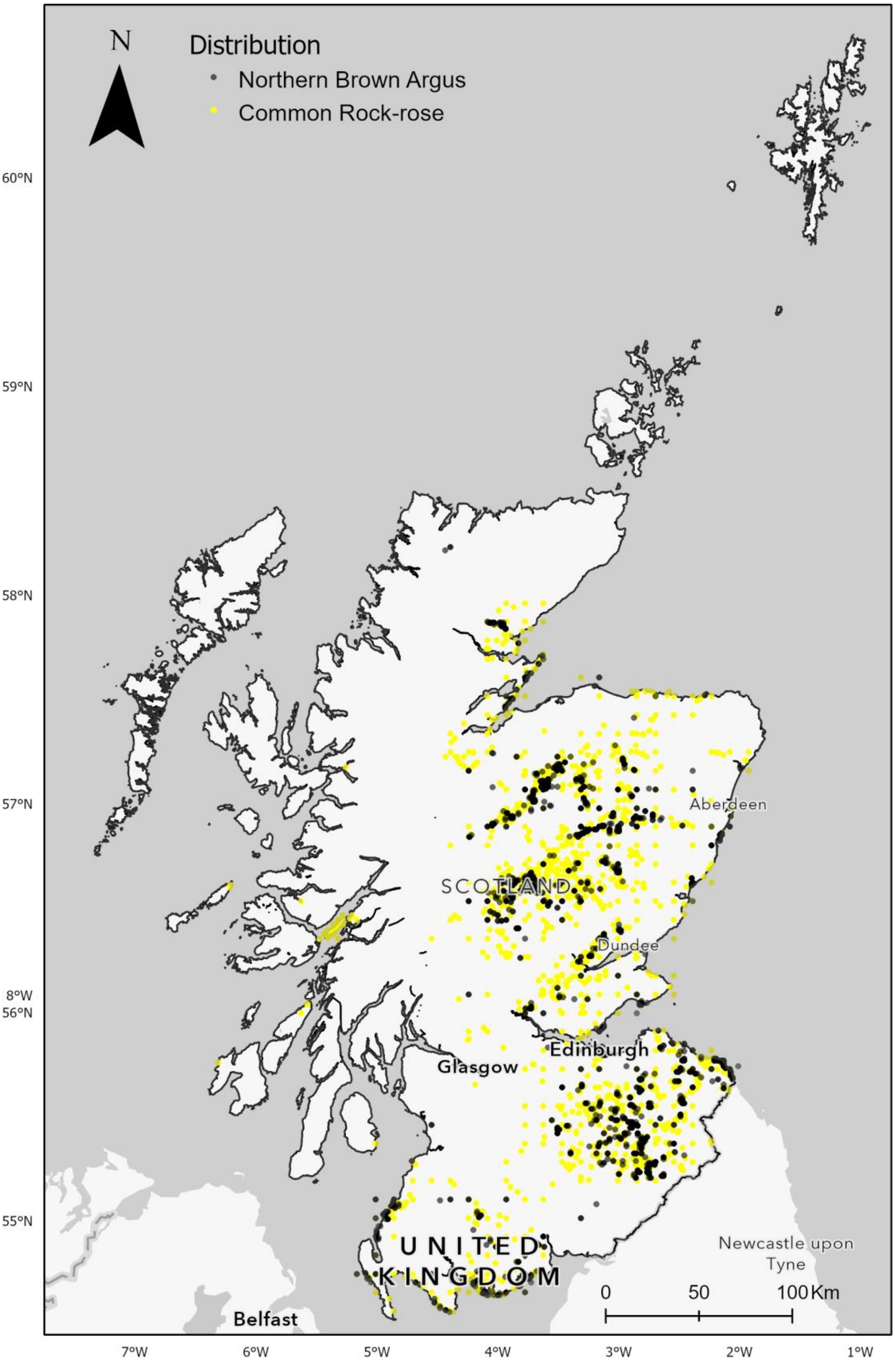


Figure 2-1. Distribution of the Northern Brown Argus Butterfly (*Aricia artaxerxes*) and its larval food plant the Common Rock-rose (*Helianthemum nummularium*) in Scotland since monitoring began. Records accessed via NBN Atlas (Appendix A-1 and Appendix A-2 respectively for citations) and displayed from a CSV to point locations in ArcGIS Pro (Esri Inc, 2020).

A. artaxerxes has the potential to move onto other classifications of SRGs (including neutral and acid) if these outlined conditions are met. However, it must be noted that the dispersal ability of this species is weak, with movements usually recorded at less than 100 m (Gallacher, 2007). Exact climatic requirements for *A. artaxerxes* are not outlined in literature. Based on its occurrence data (found in Northern European countries e.g., Norway and Sweden) and northward distribution shift in response to climate change, it can be assumed that slightly cooler temperatures are required by this species (Mallet *et al.,* 2011; Thomas *et al.,* 2006). If climate change is to continuously push these species northward, locating sites of SRGs that contain *H. nummularium* is vital, especially in Scotland where the larval food plant is still abundant. Not only this, but *H. nummularium* is a positive indicator species for calcareous SRGs, lying dormant in viable seed banks with the possibility of quick recovery in enhanced SRGs (Shellswell *et al.,* 2016). This suggests that restoring areas of SRGs could be a feasible management option for increasing *A. artaxerxes* habitat if its larval food plant establishes quickly.

2.3 Aims and Objectives

Through this case-study, the thesis will explore the methods that will allow the two disciplines of RS and CS to be merged; how and where OS practices will facilitate this; and where stakeholder and public engagement can occur throughout. This will allow the chapter research questions in section 1.9 to be answered. To meet these, the chapter research questions can be adapted to form case study specific research questions, identified to result in the creation of an OS, interdisciplinary tool:

RQ3 -> Can a habitat classification model be created to predict species-rich grasslands in Scotland and locate habitat for vulnerable species?

RQ4 -> Is currently available open-source remote sensing data able to accurately monitor species-rich grasslands and their vulnerable species?

RQ5 -> Can citizen science data validate the outputs of remote sensing models to identify species-rich grasslands for vulnerable species protection?

2.4 Thesis Development and Structure

The introduction in Chapter 1 has explored literature to identify gaps in biodiversity monitoring and consequently determine the aims of this thesis' research. I wanted to target shortcomings in biodiversity monitoring, whilst addressing the identified gaps (such as lack of discipline integration) that exacerbate these shortcomings to provide a novel method that could help to halt and reverse biodiversity loss; one of today's biggest threats to sustainability and planet functioning. The common and emerging monitoring techniques of CS and RS were regularly mentioned but rarely united and I endeavoured to provide an example of how this could be done. To combine CS and RS in a habitat monitoring approach, certain research focuses have occurred. Due to the time and labour limitations of a PhD thesis, a specific case study needed to be the focus of the application of this novel combined tool. This is where the help of Butterfly Conservation was initiated.

Collaborating with the UK's largest conservation charity would allow great research and conservation impact as well as enhance opportunities for maximum public engagement. For this relationship to be successful, specific goals that would meet Butterfly Conservation's own conservation targets were vital. This led to the development of this Chapter 2. Initial discussions with Butterfly Conservation deliberated important indicator species that are current in their initiatives such as *Aricia artaxerxes*. These indicator species supported what the literature evidenced surrounding population declines of Lepidopteran species, and more widely the disregard of invertebrates in monitoring and conservation attempts. To capture a greater range of biodiversity and utilise the capacity of Lepidoptera as indicator species, species-rich grassland habitats were more broadly targeted, which were more suitable for the RS methodologies to be explored, as well as encompassing further UK conservation targets. This would hopefully allow the transferability of the developed novel tool explored in the thesis across habitats and species.

To make the novel tool transferable at this scale, open science practices needed to be ensured in its development. Not only this, but open science would further promote discipline integration and public engagement. This led to the research associated in the third chapter, which explores the current practices implemented in biodiversity monitoring citizen science surveys in a systematic review, to explore how well they adhere to open science frameworks. This was with the aim of guiding my own citizen science survey and targeting improvements in the process. This chapter largely provides knowledge which would later be applied in the thesis through the development of my citizen science survey.

Alongside the third chapter, the fourth chapter starts the exploration of what is currently known about the locations, classifications, and conservation practices of SRGs in Scotland. By using secondary data, spatial analyses were conducted to identify SRG sites to collect environmental and remote sensing data. Part of this secondary data was gathered from Butterfly Conservation's own records, as well as data from NatureScot, who I was able to also work with on multiple occasions throughout the thesis, as they were aiming to create a comprehensive habitat map across Scotland, which would include SRGs. The environmental data collection process for this chapter was discussed with BC before the first field campaign, as these methods would be adapted for the subsequent CS survey. The data collection methods were used over two years, 2021 and 2022, where the former data was used to inform the classification of SRGs for creating and training a habitat classification model. The classification schema for identifying SRGs focused on the various known SRG types in the UK, including calcareous SRGs, which is the specific habitat for *A. artaxerxes*.

This resulting model was applied to satellite imagery in Scotland to predict areas of SRGs, which would provide the inputs for the CS survey. This was with the aims of examining if the method works not only for broader habitats, but for the specific habitats of identified target species (in this case, in relation to the discussions held between Butterfly Conservation and myself). Only the data from 2021 could be used for this as the CS survey needed to be initiated in summer 2022, whilst any further data could be collected. During this time, the CS survey was being developed with meetings with Butterfly

Conservation to finalise the participant methodology, promotion via Butterfly Conservation's social media outlets and other outreach activities, and with an ethics application submitted to the graduate school. The project was initiated on an online platform which was disseminated with the promotion, and, after receiving ethical approval, the CS survey started to run from July 2022.

Whilst the first implementation of the CS survey was running in 2022, I was out for a further three field campaigns that summer to continue to collect environmental and remote sensing data, utilising those same methodologies from chapter four. As such, the fifth chapter investigated the application of RS techniques to SRGs, where they have been little tested before, to assess if these applications may further enhance mapping of SRGs across Scotland and help identify habitat for those priority species, such as *A. artaxerxes*. Different RS sensors were used to assess the success of spatial and spectral scales on trait and species diversity predictions across SRGs. If successful, it was with hopes that this could further improve future classification mapping of SRGs or other heterogenous grassland sites. At the end of this data collection period, the results were analysed for chapter five to see how this could improve mapping attempts. The results from the CS survey were explored, which identified shortcomings in participant uptake. As such, further discussions with Butterfly Conservation were had to discuss potential reasons and solutions. This led to the adaptation of the CS survey methods that would make up chapter six. A more targeted implementation of the CS survey was decided upon to try increase participant uptake, whilst also resulting in greater stakeholder engagement. Further promotion, methodology adaptations, secondary data collection, and stakeholder collaboration occurred over the subsequent year, and in summer 2023 the CS survey ran again from May to August.

Whilst the earlier chapters in this thesis worked more independently on developing each individual approach, the sixth chapter aimed to fully merge the two disciplines and examine the results from combination of CS and RS. The sixth chapter implemented the RS outputs into the CS survey and investigated the success of the approach for habitat mapping. This research utilised citizen scientists to see how well the habitat classification model predicted SRGs and specific priority species habitat, such as the calcareous SRGs for *A. artaxerxes*. It also investigated the confidence and experience of citizen scientists in biodiversity research and how this could have affected the results the CS survey. The collected secondary data from other stakeholders allowed further analysis into the success of various participants interacting the RS data. The results of the final analysis are to be openly shared, if not already, both with the stakeholders and the wider public online via the project platform.

The thesis finished with a final seventh chapter that synthesises the previous chapters. The results are summarised, whilst the achievements and limitations of the methods tested are analysed. The wider applications and future considerations of the novel biodiversity monitoring tool are discussed, finishing with concluding remarks. The flow of information, input of data, and overlap of methods can be visualised in Figure 2-2, that leads to the final synthesis chapter.

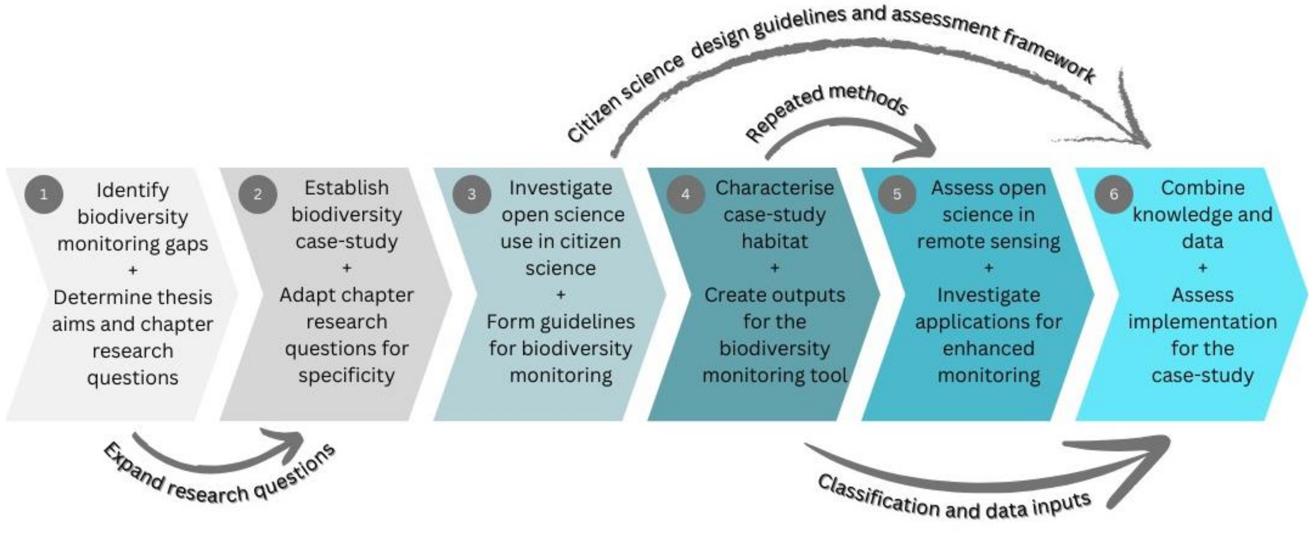


Figure 2-2. Thesis structure development.

References

Ellis, S., 2003. Habitat quality and management for the northern brown argus butterfly *Aricia artaxerxes* (Lepidoptera: Lycaenidae) in North East England. *Biological Conservation*, *113*(2), pp.285-294.

Fleishman, E. and Murphy, D.D., 2009. A realistic assessment of the indicator potential of butterflies and other charismatic taxonomic groups. *Conservation biology*, *23*(5), pp.1109-1116.

Fox, R., Dennis, E.B., Purdy, K.M., Middlebrook, I., Roy, D.B., Noble, D.G., Botham, M.S., and Bourn, N.A.D., 2023. The State of the UK's Butterflies 2022. Butterfly Conservation, Wareham, UK.

Gallacher, J., 2007. The northern brown argus in40oil4040oil40 hi40oil40mannanshireshire. *Forth Naturalist and Historian*, *30*, p.73.

Habel, J.C., Samways, M.J. and Schmitt, T., 2019. Mitigating the precipitous decline of terrestrial European insects: Requirements for a new strategy. *Biodiversity and Conservation*, *28*(6), pp.1343-1360.

Hallmann, C.A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., Stenmans, W., Müller, A., Sumser, H., Hörren, T. and Goulson, D., 2017. More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PloS one*, *12*(10), p.e0185809.

Hance, J., 2019. The great insect dying: Vanishing act in Europe and North America. *[Online]*. Accessed on 11th December 2020. Available at: https://news.mongabay.com/2019/06/the-great-insect-dying-vanishing-act-in-europe-and-north-america/

Hayhow, D.B., Eaton, M.A., Stanbury, A.J., Burns ,F., Kirby, W.B., Bailey, N., Beckmann, B., Bedford, J., Boersch-Supan, P.H., Coomber, F., Dennis, E.B., Dolman, S.J., Dunn, E., Hall, J., Harrower, C., Hatfield, J.H., Hawley, J., Haysom, K., Hughes, J., Johns, D.G., Mathews, F., McQuatters-Gollop, A., Noble, D.G., Outhwaite, C.L., Pearce-Higgins, J.W., Pescott, O.L., Powney, G.D., and Symes, N., 2019. The State of Nature 2019. *The State of Nature partnership*. Accessed 4th January 2021. Available at: https://nbn.org.uk/wp-content/uploads/2019/09/State-of-Nature-2019-UK-full-report.pdf

JNCC., 2010. *Handbook for Phase 1 Habitat Survey – a Technique for Environmental Audit*. Joint Nature Conservation Committee, Peterborough

Magnificent Meadows., n.d. Common Rock-rose *Helianthemum nummularium*. *[Online]*. Accessed 17th February 2021. Available at: http://www.magnificentmeadows.org.uk/assets/pdfs/Common_rockrose.pdf

Mallet, J., Wynne, I.R. and Thomas, C.D., 2011. Hybridisation and climate change: brown argus butterflies in Britain (Polyommatus subgenus *Aricia*). *Insect Conservation and Diversity*, *4*(3), pp.192-199.

Natural England., 2020. Climate Change Adaptation Manual (NE751). *[Online]*. Accessed 11th December 2020. Available at: http://publications.naturalengland.org.uk/publication/5679197848862720

NatureScot., 2011. Guide to species-rich grasslands. *[Online]*. Accessed 7th December 2020. Available at: https://www.nature.scot/guide-types-species-rich-grassland

O'Mara, F.P., 2012. The role of grasslands in food security and climate change. *Annals of botany*, *110*(6), pp.1263-1270.

Oostermeijer, J.G.B. and Van Swaay, C.A.M., 1998. The relationship between butterflies and environmental indicator values: a tool for conservation in a changing landscape. *Biological conservation*, *86*(3), pp.271-280.

Plantlife., 2018. Grasslands action plan. *[Online]*. Accessed 7th December 2020. Available at: https://www.plantlife.org.uk/application/files/2315/3087/2058/Grasslands_action_plan_-_Plantlife.pdf

Rodwell, J.S. ed., 1998. *British Plant Communities: Volume 3, Grasslands and Montane Communities* (Vol. 3). Cambridge University Press.

Sánchez-Bayo, F. and Wyckhuys, K.A., 2019. Worldwide decline of the entomofauna: A review of its drivers. *Biological conservation*, *232*, pp.8-27.

Scudder, G.G., 2017. The importance of insects. *Insect biodiversity: science and society*, *1*, pp.9-43.

Seibold, S., Gossner, M.M., Simons, N.K., Blüthgen, N., Müller, J., Ambarlı, D., Ammer, C., Bauhus, J., Fischer, M., Habel, J.C. and Linsenmair, K.E., 2019. Arthropod decline in grasslands and forests is associated with landscape-level drivers. *Nature*, *574*(7780), pp.671-674.

Shellswell, C.H., Chant, J.J., Jefferson, R.G., Le Bas, B., Edwards, J. and Parton, C., 2016. Restoration of existing lowland grassland – timescales to achieve favourable condition. Plantlife, Salisbury. Accessed 25th January 2021. Available at:

http://www.magnificentmeadows.org.uk/assets/pdfs/Lowland_grassland_timescales_for_recovery_advisory_note_FINAL-Design.pdf

The Wildlife Trusts., n.d. Common Rock-rose. *[Online]*. Accessed 17th February 2021. Available at: https://www.wildlifetrusts.org/wildlife-explorer/wildflowers/common-rock-rose

Thomas, C.D., Franco, A.M. and Hill, J.K., 2006. Range retractions and extinction in the face of climate warming. *Trends in Ecology & Evolution*, *21*(8), pp.415-416.

UK Biodiversity Action Plan. 2008. Priority Habitat Descriptions. BRIG (ed. Ant Maddock).

Van Swaay, C.A., Dennis, E.B., Schmucki, R., Sevilleja, C.G., Balalaikins, M., Botham, M., Bourn, N., Brereton, T., Cancela, J.P., Carlisle, B. and Chambers, P., 2019. The EU Butterfly Indicator for Grassland species: 1990-2017: Technical Report. *Butterfly Conservation Europe & ABLE/eBMS. www. butterfly-monitoring. net*.Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., Samson, R. and Wagenknecht, K., 2021. The Science of Citizen Science Evolves. *Chapter 1. in Vohland, K. et al (Eds.).(2021) The Science of Citizen Science. Springer. https://doi. org/10.1007/978-3-030-58278-4. pp 1-12.*

Walker, K.J., Stevens, P.A., Stevens, D.P., Mountford, J.O., Manchester, S.J. and Pywell, R.F., 2004. The restoration and re-creation of species-rich lowland grassland on land formerly managed for intensive agriculture in the UK. *Biological conservation*, *119*(1), pp.1-18.

Walton, P., Eaton, M., Stanbury, A., Hayhow, D., Brand, A., Brooks, S., Collin, S., Duncan, C., Dundas, C., Foster, S., Hawley, J., Kinninmonth, A., Leatham, S., Nagy-Vizitiu, A., Whyte, A., Williams, S. and Wormald K., 2019. The State of Nature Scotland 2019. *The State of Nature partnership*. Accessed 4th January 2021. Available at: https://www.nature.scot/sites/default/files/2019-10/State-of-nature-Report-2019-Scotland-full-report.pdf

Chapter 3. Do Biodiversity Monitoring Citizen Science Surveys Meet the Core Principles of Open Science Practices?

All authors (Samantha Suter, Natalie Welden, and Brian Barrett) contributed to the study conception and design. Material preparation, data collection and analysis were performed by Samantha Suter. The first draft of the manuscript was written by Samantha Suter and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Status: Published in *Environmental Monitoring and Assessment* on 12.01.23. Reformatted for thesis.

Suter, S., Barrett, B. & Welden, N. Do biodiversity monitoring citizen science surveys meet the core principles of open science practices?. *Environ Monit Assess* **195**, 295 (2023). https://doi.org/10.1007/s10661-022-10887-y

Abstract

Citizen Science (CS), as an enabler of Open Science (OS) practices, is a low-cost and accessible method for data collection in biodiversity monitoring, which can empower and educate the public both on scientific research priorities and on environmental change. Where OS increases research transparency and scientific democratisation, if properly implemented, CS should do the same. Here we present the findings of a systematic review exploring "openness" of CS in biodiversity monitoring. CS projects were scored between -1 (closed) to 1 (open) on their adherence to defined OS principles: accessible data, code, software, publication, data management plans, and preregistrations. Openness scores per principle were compared to see where OS is more frequently utilised across the research process. The relationship between interest in CS and openness within the practice was also tested. Overall, CS projects had an average open score of 0.14. There was a significant difference in open scores between OS principles ($p = <0.0001$), where "open data" was the most adhered to practice compared to the lowest scores found in relation to preregistrations. The openness of OS principles did not change in CS publications throughout publication years ($p = >0.05$ per principle). These results reveal CS is not generally "open" despite being an OS approach, with implications for how the public can interact with the research that they play an active role in contributing to. The development of systematic recommendations on where and how OS can be implemented across the research process in citizen science projects is encouraged.

Keywords: citizen science, open science, environmental monitoring, public, volunteer, data

3.1 Introduction

3.1.1 Open Science and Research Democratisation

The increasing efforts to democratise science and its outputs have resulted in the rapid development of novel research practices and resources (Mirowski, 2018; Strasser and Haklay, 2018). One coordinated approach by which we may improve the accessibility and transparency of research to readers of all backgrounds is the Open Science movement. Open Science (OS), although not a new term, may be hard to categorise under a single definition. However, it can be generally understood in its aim to increase the availability of all scientific research to general society (be that policymakers, laymen, or other researchers), or to develop a "transparent and accessible knowledge that is shared and developed through collaborative networks" (Vicente-Sáez and Martínez-Fuentes, 2018). Proponents of OS argue that this availability leads to increased reproducibility of research, as the inquiring body has access to the entire research process, including data, code, methods, analysis, and results (Taylor *et al.*, 2017).

The OS movement has become ever more important with increased public interaction with research, both through social media and through continuous news in the online era. Research affects society via its potential to inform policy, influence the economy, design technology, and effect sociality. Research is frequently funded by public investment from the taxpayer; for example, in 2017, £9 billion was spent on research and development in the UK (The Royal Society, 2019). The UK government also increased this spending by 15% in 2021, meaning the return of investment should be even greater for the public (Stokstad, 2020). However, there are numerous barriers between scientific research and the successful implementation of research findings (e.g., in policy) which may be addressed using OS practices. This has previously been demonstrated in the development of conservation policies using open data collected by the citizen science project 'eBird', turning the research into tangible achievements (Sullivan *et al.,* 2017). Practicing OS has also increased wide collaboration, as seen in the development of the Zika, Ebola, and COVID-19 vaccines, providing novel mRNA methods which now shape the future of disease response (Burgelman *et al.,* 2019; Edwin *et al.,* 2020; Pardi *et al.,* 2018).

Research's reliance on public funding and potential societal impact underlines the importance for every individual, no matter their role in society, to have access to the research that shapes their lives. Nevertheless, research is seldom available to those it influences. Indeed, access to both publications and data remain stubbornly uneven across the scientific community, and research not available to the whole scientific community is also less likely to be available for the public (Scaria and Rangarajan, 2016). However, the OS movement is not only economically appealing, due to the apparent return on investment that research may bring (e.g., in the creation of new products or industrial innovation), it also fosters trust, enhanced knowledge, and awareness of what the public are contributing towards (Grand *et al*., 2012; Houghton *et al*., 2010; OECD, 2015).

3.1.2 Open Science and Research Integrity

Applications of OS also address barriers present when seeking to determine the reliability of research. For example, lack of reproducibility in research prevents work from being verified, replicated, and expanded upon (Scaria and Rangarajan, 2016). OS can improve the quality and reliability of research through increased opportunity for peer review, method replication, and collaboration. As this can increase the robustness of the methodology, published reports may be at reduced likelihood of being retracted. Additionally, papers which had associated open data were even less likely to be retracted (Lesk et al., 2019). This has also been found with research that linked preprints, with 0.03% retracted compared to studies without preprints with a retraction rate between 0.04-0.06% (Avissar-Whiting, 2022). Where retractions are due to scientific misconduct, increased transparency in research should improve research integrity and reduce retractions overall (Marcus and Oranksy, 2014). In fact, Marcus and Oransky (2012) called for journals to have a transparency index; essentially a metric that measures processes in journal publishing, including the employment of OS practices. In addition to increasing the reliability of the publication base, the practice of OS has additional benefits to researchers; increasing the publication of null findings, increasing citations from preregistrations, improving researcher rights, and heightening collaboration leading to greater research efficiency (Clements, 2017; Franco et al., 2014; Hajjem et al., 2006; Levin et al., 2016; McKiernan et al., 2016).

3.1.3 Open Science Approaches

The practice of OS may be divided into five “advocacy schools”, including: a) public influence on, and understanding of, scientific research; b) the accessibility to both re-use raw data and retrieve published results; c) the architecture surrounding the storing and dissemination of research, d) the collaboration between different parties to increase both inputs and outputs of the research, and e) how to measure the research impact (Fecher and Friesike, 2014). Although, implementation of the aims of these OS schools will be specific to the area of the research process in question.

Various models have been proposed to highlight OS approaches (Table 3-1). The relative newness and broadness of OS, however, has meant that there are no strict guidelines to follow for practising OS. Therefore, the application across the multiple models is often quite specific. For example, Klein *et al.* (2018) produced a framework for OS in psychological research to demonstrate how, and where, to open the research process, whereas Ayris *et al.* (2018) defined 8 pillars of OS for university practice. Ayris *et al.’s* (2018) guidelines are less specific in their instructions on where OS can be implemented but instead broadly examine what it means to be practising within OS. Other models, such as Bowman and Keene’s (2018) “onion” model and de la Fuente’s (2019) “beehive” model do highlight specific practices, however, are less obvious in where these fall within the research process. Common OS practices do appear across the models which can be used to create a general framework to highlight where OS practices can be used across the research process.

Table 3-1. Published open science practices and their applications in research.

OS Practice	Application in Source	Source
FAIR Data, Research Integrity, Next Generation Metrics, Future of Scholarly Communication, Citizen Science, Education and Skills, Rewards and Initiatives, and European Open Science Cloud	University practice, research methodology	Ayris *et al.*, 2018
By author request, shared materials, shared analysis, shared data, pre-registered reports	Cross-disciplinary, research methodology	Bowman and Keene, 2018
Open notebooks, open data, open peer review, open access, open source, scientific social networks, citizen science, and open educational resources	Cross-disciplinary	de la Fuente, 2019
Data management plan, published pre-registers, materials (data and scripts) on public repository,	Psychology research process	Klein *et al.*, 2018

OS practices at the start of the research process include the formulation of data management plans (DMP) and preregistrations. DMPs are of increasing importance to many journals and funding bodies and allow the researcher to consider how they will handle, store, and share the data collected during a study. Creating a DMP allows OS to be considered at the start of a project and ensure that it can be practiced throughout the research process. (Williams *et al.*, 2017). The process of locating DMPs must be straightforward to improve efficiency and, as such, the adoption of a key sharing platform should be recognised. In the UK, many institutions encourage the use of DMPonline as a major sharing platform (Simms and Jones, 2017). Preregisters require detailing on hypotheses to be tested, methods for data collection, and the analysis to be undertaken. By describing how the data is going to be analysed it removes the possibility of changing the analysis method depending on the study findings or generating hypotheses after the results were found (known as HARKing) (Parker *et al.*, 2019). Preregisters use online platforms for sharing, for example, the largest is the Open Science Framework (Kupferschmidt, 2018).

Discoverable research outputs are the most widely practiced aspects of OS. As indicated above, open data enables research to be replicated and verified, as well as increasing collaboration. There are many platforms available to store and publish data (such as, DataONE and the Register of Research Data Repositories) where data can be located across several repositories (Michener, 2015). Similar tools can be used to publish methods and code for data analysis, but it is important to make sure that the analytical programs used are accessible to all, for example, with the use of open-source data analysis software. However, the sensitivity of data must be considered and may form a constraint on what data can be shared. Recognition of which has given rise to the term “as open as possible and as closed as necessary” (European Commission, 2016). Finally, there is the focus on open access results. Many journals have subscription fees, or one-off payments to allow access to a published paper. This reduces potential engagement with other researchers and the public (Peterson *et al.*, 2019). Access to publications is the most frequently prioritised OS practice, however, it is important to have all OS tools in place throughout the research process, otherwise, each area will lose significance. The entire research process and the outputs must be transparent to reproduce, verify, and expand upon research (Figure 3-1).

Figure 3-1. How open science can be implemented along the research process and how citizen science can achieve these.

3.1.4 Citizen Science as an Open Science Practice

Science communication and public participation are both emphasised in the practice of OS. Citizen science (CS) involves the public in the scientific research process, most commonly for data collection and analysis purposes (Cohn, 2008), thus integrating both of these elements. The collaborative aspect of CS may overcome the disconnect between the scientist, the policymaker, and the public through collaboration during the research process (Cavalier and Kennedy, 2016). In this manner, CS both enables public access to research and integrates knowledge exchange, allowing contributions from the public (Hecker *et al.*, 2018). This betters the efforts of OS at growing science democratisation through increased knowledge exchange, understanding of the scientific process, and diverse representation, with research that is more aligned with the public interest (Strasser and Hakley, 2018). The greatest benefit of CS in OS may be seen in the public advocacy school. For example, for species monitoring in Europe there are 18 citizen scientists for every 1 research scientist (Groom *et al.*, 2017). It is important to harness this enthusiasm from the public to generate idea creation and democratic governance (Storksdieck *et al.*, 2016).

An important contribution of CS to OS practices should be the increased availability of the data to potential collaborators who are not involved directly with the research. Many of the largest CS projects are in biodiversity monitoring (Kullenberg and Kasperowski, 2016). Indeed, such collaboration is essential in tackling the current biodiversity crisis (Costello *et al.*, 2015). Both CS, and more broadly OS, enable new partnerships to address the gaps in global biodiversity monitoring. Firstly, OS can do this by developing platforms, such as the Group on Earth Observations Biodiversity Observation Network or Collect Earth, where common data and information is shared (Pettorelli *et al.*, 2014). Secondly, CS can then combine datasets across a larger scale, for example where the Euro Bird Portal in Europe and the Second Southern African Bird Atlas Project in Africa pooled data across multiple countries to compile one large dataset (Amano *et al.*, 2016).

As indicated above, OS principles require that research should be reproducible and that data be accessible, and it may be assumed that CS adheres to these principles. However, Groom *et al.* (2017) found that CS data scored the second lowest on an open data index concerning biodiversity observations. Reasons for this included licensing restrictions, surrounding landowner permissions, concerns from data holders on data sharing, and funding disincentives (Groom *et al.*, 2017). However, open CS data has the potential to provide large amounts of data which may otherwise not be achieved; this was seen in a water monitoring scheme in the US, where the CS data was made open, forming more than 50% of observations (Poisson *et al.*, 2020). If CS data remains closed this can restrict further research on a topic by hindering collaboration. To overcome this issue, the reasons for reduced data sharing need to be addressed. For example, how can sensitive data be protected if it is openly available? Once these issues have been addressed the benefits of OS can be harnessed.

3.2 Aims and Objectives

In this chapter, I am addressing the thesis RQ2: To what extent do citizen science studies of biodiversity demonstrate the principles of open science? For CS to fulfil its potential as a core practice in OS, CS projects must adhere to the full range of OS principles. Although previous studies have indicated the inaccessibility of CS data, CS projects have not been assessed in all areas of the research process. As such, this study aims to systematically review biodiversity monitoring CS projects to determine whether they meet the core principles of OS. The study focuses on the research process from planning through to result reporting (step 2 – 5, Figure 3-1). Collaboration in step 1 was not included in the following analysis, due to its multi-layered nature and to simplify the scoring system across the principles. The specific research questions identified to meet RQ2 were:

i) How open are environmental citizen science projects across the entire research process?
ii) Which aspect of the research process is most open and widely adhered to?

iii) Has openness in OS principles in published environmental citizen science projects changed over time?

3.3 Methods

3.3.1 Selection of Studies

A systematic review was undertaken to identify publications arising from CS projects on biodiversity. The key words “CS” and “biodiversity monitoring” were used to search Web of Science. This database was chosen due to its common use in systematic reviews, specifically surrounding biodiversity conservation research, and the large number of results generated in the initial search (Boice, 2019). Results were initially filtered to include only journal article submissions before 2022.

Subsequently, the results were exported to EndNote and duplicates were removed, before being sorted by title and abstract to exclude review papers, book entries, irrelevant studies (e.g., studies which did not consider CS or biodiversity monitoring), studies using secondary data or data not based on the authors’ own CS projects, or studies that were based on development of certain aspects of a CS project (for example, an app). Where it was unclear whether these criteria were met, the publication was carried over into the subsequent sorting stage. The methods sections of each paper were then reviewed and those papers which did not have enough relevant information to meet the criteria above were excluded. The number of papers returned at each stage of this sorting process can be seen in Figure 3-2.

Due to a large level of incomplete information (especially regarding data management plans and preregisters), the authors of the 153 papers identified from the methods sorting were all contacted regarding missing information. Authors were given a month to respond to the emails before being excluded. Projects by authors that gave complete answers were included for analysis (Figure 3-2).

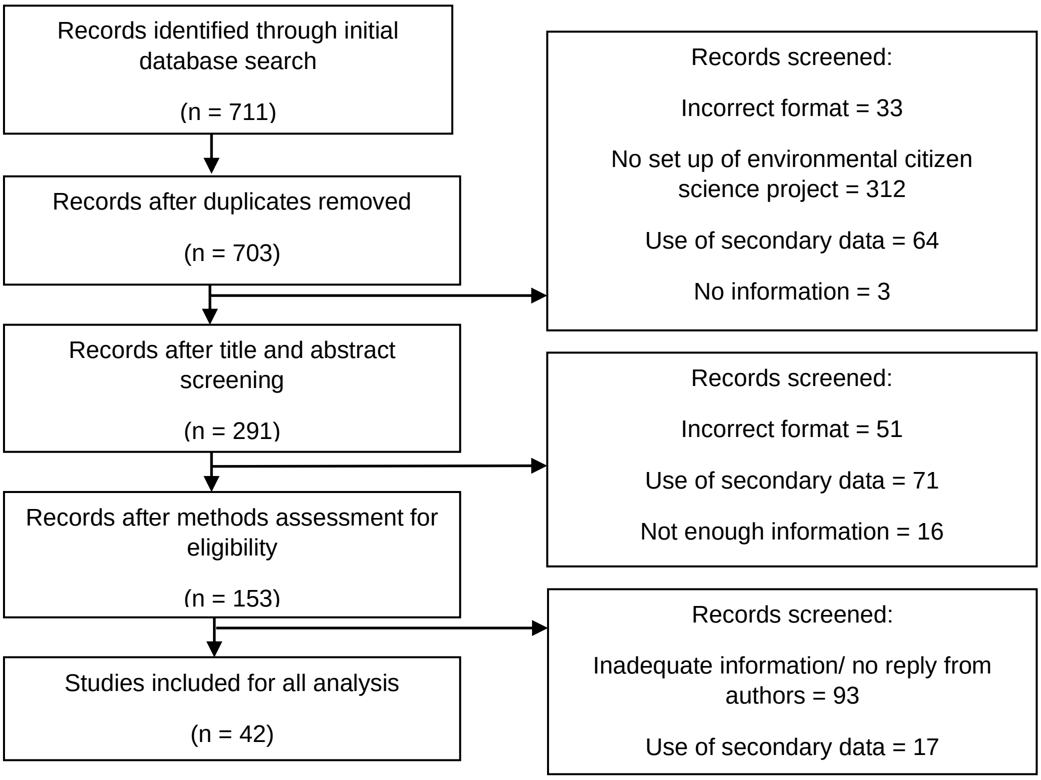


Figure 3-2. PRISMA diagram of the selection and exclusion process of included papers in the systematic review.

3.3.2 Open Science Criteria

The CS projects identified were assessed based on OS criteria. These included whether studies had associated DMPs, preregisters, available data and code, free software for analysis, and open findings available to everyone. Projects were assessed to see how many of the OS principles were met out of the six. Projects were then scored based on how open each principle was; closed was given a score of -1, partially open was given a score of 0.5, and fully open was given a score of 1 (Table 3-2). Where the OS principle was not applicable (for example, where projects used software that does not require code, or where simple descriptive statistics were used which did not require specific packages) these were left blank.

Table 3-2. Open science criteria used to rate the openness of biodiversity monitoring citizen science projects between -1 and 1. Each criterion was scored from -1 (closed), 0.5 (partially open), and 1 (open). Where information was not applicable to the project these were left blank.

OS Criteria	Closed (-1)	Partially Open (0.5)	Open (1)
Data Management Plan	None	Accessible by request/internal	Discoverable (e.g., DMPonline/DMPtool)/attached
Preregistration	None	Accessible by request/internal	Discoverable (e.g., Open Science Framework)/attached
Open Data	None	Accessible by request/internal	Attached/online repository
Open Code	None	Accessible by request/internal	Attached/online repository
Open Software	Subscription software	Transferable	Free software
Open Access	Behind paywalls/subscriptions	Partial access	Open access journals/websites for results

3.3.3 Data Analysis

All statistical analysis was conducted in R (v3.6.3). To answer how open CS projects are across the entire research process, the scores for each project (n = 42) were averaged to give a value between -1 to 1, with higher values signifying more open projects. Mean score was selected rather than total score, as not all categories were relevant to each paper (for example, open code), and the use of totals artificially penalised these studies. A final average was calculated across all projects to give an average "openness" score along the research process. The scores of each OS principle per project were used to investigate which aspect of the research process is the most open. Assumptions for normality could not be met and a non-parametric Kruskal Wallis test was used to evaluate whether any of the OS principles were more commonly applied. Ordinal logistic regression models were used to investigate the change in openness of OS principles in CS projects over time. Here, each principle per projects' openness category was analysed across the years that the projects were published. Year was determined as project publication years. All significance levels were set to 0.05. The R coding script can be found in Appendix B-1.

3.4 Results

3.4.1 Average Openness Across the Research Process

Analysis was conducted on 42 biodiversity CS projects published between 2005 and 2021 (see supplementary material, Appendix B-2), with an average openness score of 0.14 (minimum -0.67, maximum 0.9) across all projects for all OS principles combined. However, the number of papers (n = 153) identified during the review process was initially much higher, with many papers discounted due to missing information. For example, only 36.6% of the original 153 papers had a data availability statement and/or a supplementary material section. However, this did not always detail whether either the full data and/or code or preregisters and/or data management plans were attached or completed. The authors of these papers were contacted in the event that they should wish to provide further detail regarding their studies. All original 153 authors were contacted; 69 responded, however, only 42 projects provided complete information. The only OS practice that could be scored without correspondence was the openness of findings (i.e., was the paper published under an open access license or not). Of the original 153 papers, 152 papers had to be contacted regarding preregisters.

3.4.2 Adherence to Open Science Principles

There was a large variation in adherence to OS principles across projects (Table 3-3). The distribution of openness scores across projects, per OS principle can be seen in Figure 3-3. The most regularly adhered to OS principle was Open Data (69%), followed by Open Access where 64% of papers had open access publishing. The percentage of projects that had fully open software was 58%, whereas 35% of projects had fully open code, 12% had fully open DMPs, and just 7% had fully open Preregisters.

Table 3-3. The number of citizen science projects distributed by their openness across open science principles.

OS Principle	Number of Projects per Principle (after NA removal)	Number of Projects by Openness Score		
		-1 (Closed)	0.5 (Partially Open)	1 (Open)
Open Access	42	15 (35.7%)	0 (0%)	27 (64.2%)
Open Code	26	9 (34.6%)	8 (30.8%)	9 (34.6%)
Open Data	42	5 (11.9%)	8 (19%)	29 (69%)
Open DMP	41	26 (63.4%)	10 (24.4%)	5 (12.2%)
Open Preregistration	41	27 (65.9%)	11 (26.8%)	3 (7.3%)
Open Software	33	5 (15.2%)	9 (27.3%)	19 (57.6%)

Mean openness scores for each principle across the 42 projects were calculated after removal of non-applicable categories. The OS principle which is most frequently employed is the use of open data with a mean score of 0.67 (± 0.65 SD) across projects. The OS principle that is least employed is the use of preregisters with a mean score of -0.45 (± 0.78 SD) across projects (Figure 3-3).

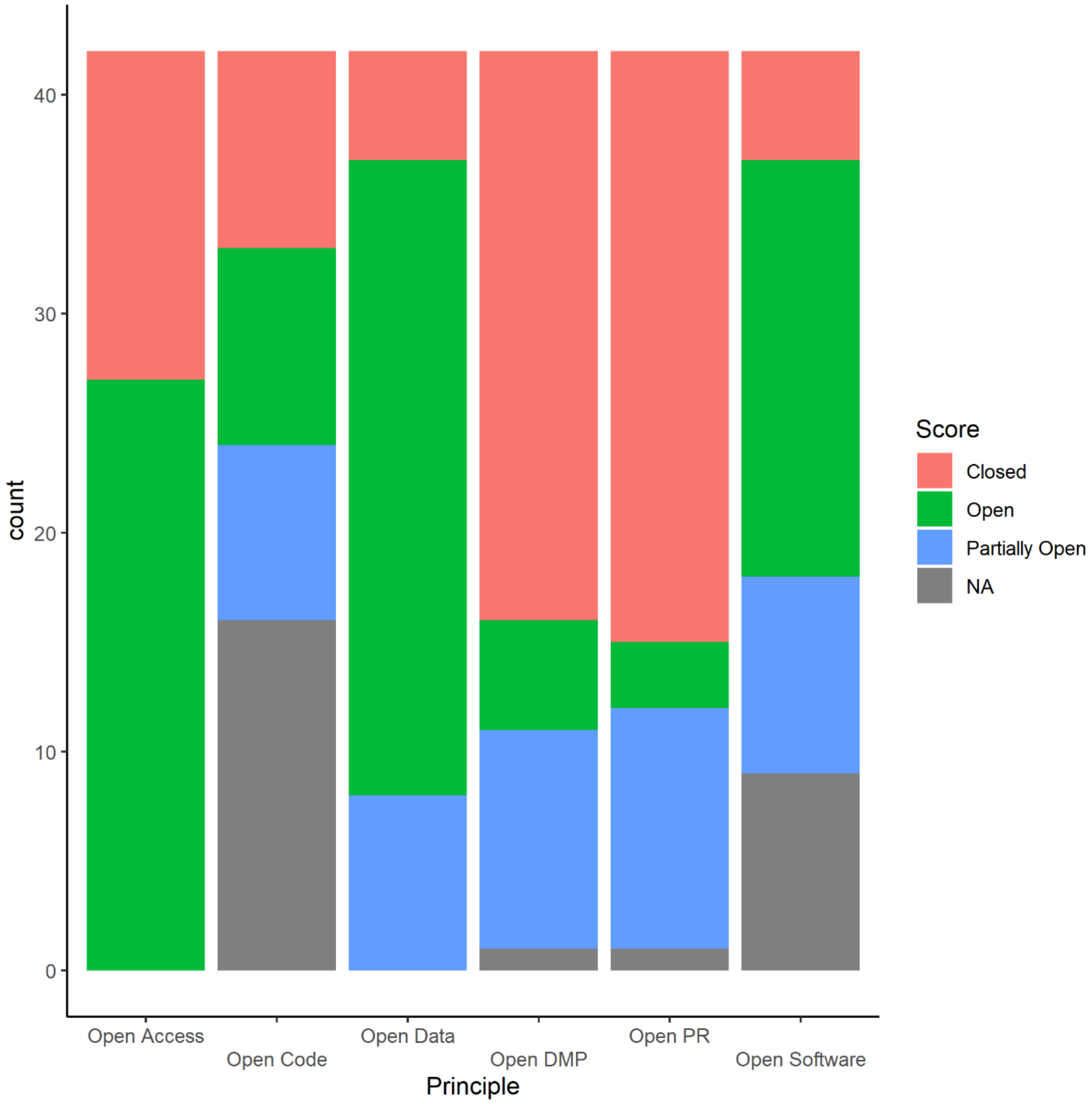


Figure 3-3. Distribution of openness scores across biodiversity monitoring citizen science projects per open science principle

The results of the Kruskal-Wallis test showed that the difference in openness scores between OS principles was significant (p = <0.0001, df = 5, F = 60.002). A post-hoc Dunn's test with Bonferroni correction showed there was a significant difference in scores between Open DMP and Open Access principles (p = 0.0006), Open Preregistrations and Open Access principles (p = 0.0001), Open DMP and Open Data principles (p = <0.0001), Open Preregistrations and Open Data (p = <0.0001), Open Software and Open DMP principles (p = 0.0001), and Open Software and Open Preregistration principles (p = <0.0001).

3.4.3 Changes to Adherence of Open Science Over Time

The number of CS projects in this review that were found in publications increased from 1 project in 2005 to 11 projects in 2021. Over this period, the average openness score did not significantly change in published CS surveys through publication years (z = 0.171, p = 0.864). It was also investigated whether openness in CS was influenced by the rise in CS projects in publications, as a proxy for increased awareness of OS. The results show there is not a significant correlation in project openness across publication year and the number of CS projects published each year ($z = -0.043$, $p = 0.966$) (Figure 3-4). Furthermore, none of the OS principles present significant changes in openness with the rise in CS projects in publications over time. The results from the ordinal logistic regression show that the openness of each principle did not significantly change over time of the published CS projects: open access ($p = 0.636$), open code ($p = 0.581$), open data ($p = 0.81$), open DMP ($p = 0.667$), open PR ($p = 0.437$), open software ($p = 0.0962$).

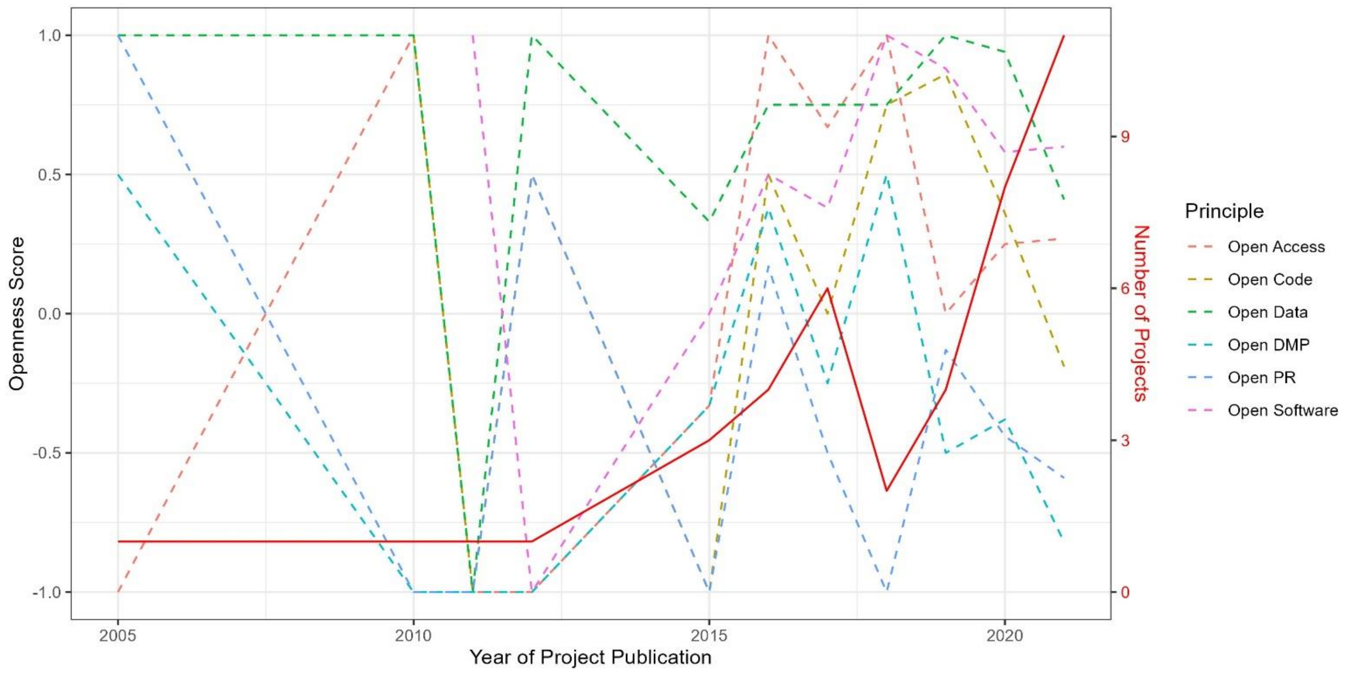


Figure 3-4. Open science principle's yearly average open scores by the number of citizen science projects per year between 2005 – 2021.

3.5 Discussion

3.5.1 Implications of Results

The application of OS principles in CS projects have not been widely investigated beyond the scope of data assessments (Borda *et al.*, 2020; Williams *et al.*, 2018). This makes comparisons among research difficult and suggests why this review is important in identifying where best practice methods are underutilised. OS practices are difficult to outline due to the varying nature of research questions and disciplines, suggesting that what is applicable to one research question may not be suitable for another. Subsequently, researchers may have different opinions on what is considered an OS practice because of the lack of consistent guidance.

The results presented above indicate a lack of consistency in the number and extent to which OS principles are commonly implemented in CS research. Open access and open data were far more common in the research process compared to other practices, with these principles also having the highest mean openness scores found in this review. Bowser *et al.* (2020) found even higher levels of discoverable data where 75% of 36 projects (across a range of scientific disciplines) made the CS raw data available by some means. However, the remaining lack of access to information (publications and data) was still noted as one of the bigger downfalls in CS projects in comparison to others, for example data quality measures. This is supported by Groom *et al.*, (2017) who found that CS data scored the second lowest on an open data index concerning biodiversity observations. OS practices (specifically around open access and open data) in biodiversity conservation have long been called for (Fonseca and Benson, 2003; Gaikwad and Chavan, 2006; Mose *et al.*, 2018). However, although open data and open access practices may score higher than other assessment criteria in reviews, they are by no means practiced on a large scale. Preregistrations and DMPs were used the least frequently. Very few studies have investigated the use of these within CS projects. One report based on CS projects across disciplines found that DMPs were implemented in 60% of projects (although it is unclear if these were publicly available) and 38% had raw open data (Schade and Tsinaraki, 2016).

The results also show that there were no significant changes in the openness scores of OS principles in published CS projects across time, nor of average project openness scores as the number of CS projects in publications increased over time either. It was theorised that a greater number of CS projects in publications could contribute to the discourse around OS and in turn result in increased openness across projects or implementation of the OS principles along the research process. However, it appears that an increase in the number of CS projects is not enough to do this. It must be noted that there are likely more effective tools that may facilitate the use of OS along the research process, for example greater discussion of OS itself, however, this could not be analysed in the study here. The non-significant changes of openness scores in each individual principle across time also suggest that either there have not been improvements to facilitate open practice within these principles or that the awareness has not increased to the extent to allow this to occur.

Although the findings in the other studies outlined above are not directly comparable to our results due to differences in study identification methods and analysis, together they do show there are still large gaps in the utilisation of OS principles. The drivers for these findings appear largely to be what is common versus what is not. Through correspondence with authors, it was noted that many researchers were unaware of certain OS practices, predominantly preregisters, making it impossible for them to be utilised. This may be the result of the lack of common practices, education, or guidelines surrounding the OS process. Similar issues were observed in relation to DMPs. DMPs can be made readily accessible by platforms, such as on DMPOnline and preregisters can be published on the OS Framework but less than 1% of papers in this review utilised these resources. Lack of guidance across scientific organisations, journals, or funders can result in fewer incentives or reduce researcher ability to contribute to OS. Allen and Mehler (2019) implied that this is likely to improve with time and greater interest and investment in OS, but this has not been assessed. Therefore, it can be suggested that the absence of an OS practice may not be an active choice researchers make but simply lack of awareness surrounding the movement.

Another common theme that appeared to limit the implementation of OS practices was the lack of funding (pers. Comms). Unfortunately, OS practices often come with associated costs, for example that of open access publishing. Many researchers are deterred by this if they do not have the resources (particularly seen in the Global South) meaning that participation in OS is not feasible or possible for all (Fontúrbel and Vizentin-Bugoni, 2021; Nabyonga—Orem *et al.,* 2020). Conversely, access to funding, or lack thereof, could explain the relatively high proportion of studies which utilised open software, identified as the third most adhered to OS principle, as open software benefits the researcher in that it is free for them to use. This review discovered that where open data or code was used, these were found in public repositories, such as GBIF, iNaturalist, github, the project's website (if applicable), or included in supplementary materials. Open software largely included R and QGIS. However, although there is software available for OS (e.g., storage, validation, and dissemination sites), many are still in the process of being created or are comparatively unknown.

It must also be noted that increased research openness may lead to increased competition. OS processes are sometimes considered more time-consuming due to the preregistration and documentation procedures involved, creating concern around reduced publications or the prospect of being "scooped" by other authors (Allen and Mehler, 2019). However, this may be counteracted by the possibility of increased citations from the open-source documents. and time reduction from collaboration opportunities. Additionally, if copyright laws that are associated with open access publishing are applied to all stages of data sharing, then the possibility of researcher's work being taken is minimised (Levin *et al.*, 2016).

Funding disincentives, time commitments, and lack of awareness imply that the widespread implementation of OS remains a low priority in biodiversity research, which is supported by this review. Previous research has indicated a rise in CS since the 20th century, with the largest application in ecological monitoring; attributed to many reasons including the low-cost, efficient methods it provides, or increased interest from the public (Kelemen-Finan *et al.,* 2018; Kullenberg and Kasperowski, 2016; Toerpe, 2013). However, the increased interest in CS has not translated to a greater adherence to OS principles within the field. Potentially, this could also be related to a lack of want within the scientific field where OS is not an accepted practice. Although this result cannot be compared to other studies it is not what was initially expected.

In this era of "big data", it was hypothesised that project openness may have increased in practice, as CS itself can facilitate OS (Bezjak *et al.,* 2018). Nevertheless, a number of meaningful restrictions on data availability remain, for example, both ethical and GDPR concerns where working with human participants makes the ethics around sharing data difficult (Suman and Pierce, 2018). As the projects are biodiversity based, there are concerns regarding the potential unintended secondary effects of data sharing, e.g., highlighting locations of rare and endangered species (Ganzevoort *et al.,* 2017). This does not mean that all principles of OS should be disregarded and justifications for certain closedness in projects should be made apparent. The results here support the notion that scientific research is often based on quantity not quality, where the number of publications and citations are seen as a measure of success (Fire and Guestrin, 2019). This is further evident from the findings that even retracted papers are still heavily cited and not removed if cited before a retraction occurs (Bolland *et al.,* 2022). One reason for this is that the retraction status of a paper is usually unknown to authors (Teixeira da Silva and Bornemann-Cimenti, 2017). Methods have been proposed where journals use software that detects plagiarism (already utilised by some) and pairs this with retraction databases, such as Retraction Watch, to increase retraction clarity (Consentino and Veríssimo, 2016). As such, OS has the potential to increase the transparency of retractions and heighten scientific credibility overall. This is especially the case when meeting the expectations of OS guidelines i.e., not take advantage of open practices such as in the publication of preprints without any peer review, which was seen during the COVID-19 pandemic (Ryzmski *et al.,* 2020). It is with hopes that OS can reduce the impact of "infodemics" and reduce scientific misinformation, not contribute to it (Pool *et al.,* 2021).

3.5.2 Limitations

This review focused on patterns of openness in biodiversity monitoring CS projects. While over 150 papers were identified in our literature review, not all the eligible studies could be included in our analysis due to incomplete project information/no response from authors. Additionally, due to time considerations, the influence of project structure on the resulting project openness and the factors which enable or are otherwise responsible for successful incorporation of OS principles were not explored. Future research may focus on these areas to target suggestions more specifically. It must also be noted that the purpose of this review is not to criticise the assessed projects for their adherence, or lack of, to OS practices, but to highlight areas for improvements that may be made in the field at large.

The scope of the included projects was wide; some projects aimed to create a CS project to assess this method as a valid tool for biodiversity monitoring, other projects used CS projects as a secondary goal, primarily focusing on analysis of the ecological data collected. It would be reasonable to assume where the development of a CS tool for biodiversity monitoring was not a primary goal of the research, OS itself would not have been a main consideration. Included papers also comprised a large geographic area with projects based on every continent. There is the possibility that the practice of OS does not translate globally, especially with a lack of common guidelines, language barriers, and issues regarding access to OS tools (technological challenges) and funding.

When evaluating the changes in adherence to OS over time, year was determined as the year the CS project was published, rather than conceptualised. It is difficult to analyse changes in OS implementation within a project, that may have occurred throughout the length of a project's research, for projects that ran longer than a year or were not published the same year as their creation. It is uncertain, therefore, as to whether research decisions were made that were unchangeable in the initial phases of a project or if decisions were influenced by the external hinderances within practicing OS itself. Project duration also varied from one day biodiversity counts to multi-decade and ongoing projects. It was not in the scope of this systematic review to determine whether project duration influenced adherence to OS practices. There is potential for projects (particularly ongoing ones) to move towards OS approaches. Even completed projects may still make previous data and code open, for example. However, this relies on a greater change in mindset and acceptance of new research practice as beneficial for scientific research and public engagement in science. Future research may look at how project duration or how time of project creation versus time of publication may impact the openness within CS.

Finally, it is acknowledged that the final literature search returned 153 papers but full responses were only able to be gathered from 42 of these. The results are, therefore, constricted to these papers and further research could investigate the practice of open science across more biodiversity monitoring studies. Similarly, these studies are based on those that are published and may be constrained by policies within the academic publishing realm. Biodiversity monitoring citizen science initiatives outside of this context may well find further freedom to practice open science; another area for future exploration.

3.5.3 The Future of Open Science

The successful application of OS principles is dependent on the existence and visibility of appropriate tools in combination with standard guidelines. These necessary tools for OS are mainly in the form of online repositories and databases to store data, social media platforms for sharing data, and free access journals to present research outcomes (Neylon and Wu, 2009). Although such tools are now more readily available, there are still barriers which arise because of different policies across journals, funders, and governing parties. However, due to the vast expanse that scientific research covers and the variety within scientific processes flexibility is required when applying practices across disciplines. Universal OS practices may not apply to all stages of research and data management, requiring different tools. It may be more advantageous to actively encourage OS through policies, institutions, and funding bodies, whilst allowing the researcher to justify the use of OS in their research (Levin et al, 2016).

The diversity of OS methods may be intimidating to the researcher but increasing awareness and uptake of such practices will make the process more commonplace, with the potential for OS courses to be undertaken (Toelch and Ostwald, 2018). Altering researcher cultures towards the practice of OS is often noted as the most difficult task when trying to make OS the norm. OS workshops focusing on approaches to OS and why it should be practiced should be made available where applicable to breed an understanding of its importance (Ignat and Ayris, 2020).

What is clear is that the lack of widely implemented OS is frequently not the fault or choice of a researcher. It is more commonly a universal shortcoming, where funding and financial security provide the incentives to fulfil (or not) OS criteria within biodiversity CS and broader scientific research. However, the benefits of OS far outweighing financial savings. Increased collaboration as a guaranteed result of OS is more than likely to provide return of investment at a greater level than initial capital projections may predict. As we are at a crucial point in world history, with mass extinction threatening ecosystem functioning and human survival, it seems that investment in OS should be non-debatable for the success of biodiversity research and the scientific process.

3.5.4 Guidelines and Recommendations for Open Science in Citizen Science Projects

The results of this study highlight areas where OS practices can be improved in CS projects. It must be noted that OS should not be prescriptive but suggestive, implementing practices where applicable and necessary. As such, OS practices that should be encouraged where applicable include:

1) When creating a CS project, a pre-register should be made publicly available detailing the aims, methods, analysis, and dissemination of results intended at the start of the project. Where possible, the use of the Open Science Framework for publishing should be engaged if the target journal does not offer this service.
2) A data management plan should be created detailing the collection, analysis, and storage of data using online creation tools such as DMPOnline if your organisation does not offer an alternative. These should be available on the project website (if applicable), as supplementary material, or published in the intended journal, as well as on DMPOnline.
3) Data and, where relevant, code should be made publicly available as supplementary material or on repositories such as GitHub/GBIF etc. and linked within the published journal article/on project websites (if applicable). Nondisclosure of sensitive data (where ethics and anonymisation cannot be instigated) should be justified within a data statement.
4) Where possible/applicable, projects should consider the use of open software for replicability for researchers, as well as useability for the public.
5) Project results should be published under an open access license and on project websites, freely available to the public.

3.6 Conclusions

CS is considered an OS practice that is implemented most often in biodiversity monitoring. Here, CS can be both a result of OS and an instigator. However, previous studies show that CS projects often do not adhere to OS practices, hindering its potential to reach the goals of OS. The results of this review show that although interest in CS has increased in biodiversity monitoring over time, the openness of such projects has not risen with this. Although principles of OS need improvement, the areas that need addressing specifically appear to be around the use of preregisters and data management plans, which should be implemented at the start of a project. Guidelines are set out to advise projects on how they can initiate more OS principles within CS projects, whilst OS is actively encouraged on a larger scale through instigation within organisations, institutions, and governments allowing scientific research to become more comprehensible, collaborative, and transparent.

References

Allen, C. and Mehler, D.M., 2019. Open science challenges, benefits and tips in early career and beyond. *PLoS Biology*, *17*(5), p.e3000246.

Amano, T., Lamming, J.D. and Sutherland, W.J., 2016. Spatial gaps in global biodiversity information and the role of citizen science. *BioScience*, *66*(5), pp.393-400.

Avissar-Whiting, M., 2022. Downstream retraction of preprinted research in the life and medical sciences. *PLoS One, 17*(5), e0267971.

Ayris, P., Bernal, I., Cavalli, V., Dorch, B., Frey, J., Hallik, M., Hormia-Poutanen, K., Labastida i Juan, I., MacColl, J., Ponsati Obiols, A. and Sacchi, S., 2018. LIBER Open science roadmap. Zenodo. https://doi.org/10.5281/zenodo.1303002

Bezjak, S., Clyburne-Sherin, A., Conzett, P., Fernandes, P., Görögh, E., Helbig, K., Kramer, B., Labastida, I., Niemeyer, K., Psomopoulos, F., Ross-Hellauer, T., Schneider, R., Tennant, J., Verbakel, E., Brinken, H. and Heller, L., 2018. Open Science Training Handbook. 10.5281/zenodo.1212496. Available at: http://book.fosteropenscience.eu/. Accessed: 12th December 2022.

Boice, J., 2019. An exploration of systematic review publication trends in conservation biology journals. *Issues in Science and Technology Librarianship*, (91).

Bolland, M.J., Grey, A. and Avenell, A., 2022. Citation of retracted publications: A challenging problem. *Accountability in Research*, *29*(1), pp.18-25.

Borda, A., Gray, K. and Fu, Y., 2020. Research data management in health and biomedical citizen science: practices and prospects. *JAMIA Open*, *3*(1), pp.113-125.

Bowman, N. D., & Keene, J. R., 2018. A layered framework for considering open science practices. *Communication Research Reports, 35*(4), 363–372.

Bowser, A., Cooper, C., De Sherbinin, A., Wiggins, A., Brenton, P., Chuang, T.R., Faustman, E., Haklay, M. and Meloche, M., 2020. Still in need of norms: the state of the data in citizen science. *Citizen Science: Theory and Practice*, *5*(1).

Burgelman, J.C., Pascu, C., Szkuta, K., Von Schomberg, R., Karalopoulos, A., Repanas, K. and Schouppe, M., 2019. Open science, open data, and open scholarship: European policies to make science fit for the twenty-first century. *Frontiers in Big Data*, *2*, p.43.

Cavalier, D. and Kennedy, E.B. eds., 2016. *The rightful place of science: Citizen science*. Consortium for Science, Policy & Outcomes.

Clements, J.C., 2017. Open access articles receive more citations in hybrid marine ecology journals. *Facets*, *2*(1), pp.1-14.

Cohn, J.P., 2008. Citizen science: Can volunteers do real research?. *BioScience*, *58*(3), pp.192-197.

Cosentino, A.M. and Veríssimo, D., 2016. Ending the citation of retracted papers. *Conservation Biology*, *30*(3), pp.676-678.

Costello, M.J., Vanhoorne, B. and Appeltans, W., 2015. Conservation of biodiversity through taxonomy, data publication, and collaborative infrastructures. *Conservation Biology*, *29*(4), pp.1094-1099.

de la Fuente, G. B., 2019. What is open science: Introduction. [Online]. Retrieved July 5, 2022, from https://www.fosteropenscience.eu/content/what-open-science-introduction

Edwin, G.T., Klug, D.M. and Todd, M.H., 2020. Open science approaches to COVID-19. *F1000Research*, *9*(1043), p.1043.

European Commission, 2016. Directorate-General for Research & Innovation. *H2020 Programme Guidelines on FAIR Data Management in Horizon 2020*.

Fecher, B. and Friesike, S., 2014. Open science: one term, five schools of thought. In *Opening Science* (pp. 17-47). Springer, Cham.

Fire, M. and Guestrin, C., 2019. Over-optimization of academic publishing metrics: observing Goodhart's Law in action. *GigaScience*, *8*(6), p.giz053.

Fonseca, G. and Benson, P.J., 2003. Biodiversity conservation demands open access. *PLoS Biology*, *1*(2), p.e46.

Fontúrbel, F.E. and Vizentin-Bugoni, J., 2021. A paywall coming down, another being erected: Open access article processing charges (APC) may prevent some researchers from publishing in leading journals. *Bulletin of the Ecological Society of America*, *102*(1), p.e01791

Franco, A., Malhotra, N. and Simonovits, G., 2014. Publication bias in the social sciences: Unlocking the file drawer. *Science*, *345*(6203), pp.1502-1505.

Gaikwad, J. and Chavan, V., 2006. Open access and biodiversity conservation: challenges and potentials for the developing world. *Data Science Journal*, *5*, pp.1-17.

Ganzevoort, W., van den Born, R.J., Halffman, W. and Turnhout, S., 2017. Sharing biodiversity data: citizen scientists' concerns and motivations. *Biodiversity and Conservation*, *26*(12), pp.2821-2837.

Grand, A., Wilkinson, C., Bultitude, K. and Winfield, A.F., 2012. Open science: a new "trust technology"?. *Science Communication*, *34*(5), pp.679-689.

Groom, Q., Weatherdon, L. and Geijzendorffer, I.R., 2017. Is citizen science an open science in the case of biodiversity observations?. *Journal of Applied Ecology*, *54*(2), pp.612-617.

Hajjem, C., Harnad, S. and Gingras, Y., 2006. Ten-year cross-disciplinary comparison of the growth of open access and how it increases research citation impact. *arXiv preprint cs/0606079*.

Hecker, S., Haklay, M., Bowser, A., Makuch, Z., Vogel, J. and Bonn, A., 2018. Innovation in open science, society and policy–setting the agenda for citizen science. *Citizen science: Innovation in Open Science, Society and Policy. UCL Press, London, UK*, pp.1-23.

Houghton, J., Rasmussen, B. and Sheehan, P., 2010. Economic and social returns on investment in open archiving publicly funded research outputs.

Ignat, T., & Ayris, P., 2021. Built to last! Embedding open science principles and practice into European universities. *Insights: The UKSG Journal, 33*(1).

Kelemen-Finan, J., Scheuch, M., & Winter, S., 2018. Contributions from citizen science to science education: an examination of a biodiversity citizen science project with schools in Central Europe. *International Journal of Science Education, 40*(17), 2078–2098.

Klein, O., Hardwicke, T.E., Aust, F., Breuer, J., Danielsson, H., Hofelich Mohr, A., IJzerman, H., Nilsonne, G., Vanpaemel, W. and Frank, M.C., 2018. A practical guide for transparency in psychological science. *Collabra: Psychology*, *4*(1), pp.1-15.

Kullenberg, C. and Kasperowski, D., 2016. What is citizen science?–A scientometric meta-analysis. *PloS One*, *11*(1), p.e0147152.

Kupferschmidt, K., 2018. More and more scientists are preregistering their studies. Should you. *Science*, *10*

Lesk, M., Mattern, J.B. and Sandy, H.M., 2019, March. Are papers with open data more credible? An analysis of open data availability in retracted PLoS articles. In *International Conference on Information* (pp. 154-161). Springer, Cham.

Levin, N., Leonelli, S., Weckowska, D., Castle, D. and Dupré, J., 2016. How do scientists define openness? Exploring the relationship between open science policies and research practice. *Bulletin of Science, Technology & Society*, *36*(2), pp.128-141.

Marcus, A. and Oransky, I., 2012. Bring on the transparency index. *Scientist*, *26*(8), pp.24-24.

Marcus, A. and Oransky, I., 2014. What studies of retractions tell us. *Journal of Microbiology & Biology Education*, *15*(2), pp.151-154.

McKiernan, E.C., Bourne, P.E., Brown, C.T., Buck, S., Kenall, A., Lin, J., McDougall, D., Nosek, B.A., Ram, K., Soderberg, C.K. and Spies, J.R., 2016. Point of view: How open science helps researchers succeed. *Elife*, *5*, p.e16800.

Michener, W.K., 2015. Ecological data sharing. *Ecological Informatics*, *29*, pp.33-44.

Mirowski, P., 2018. The future (s) of open science. *Social Studies of Science*, *48*(2), pp.171-203.

Mose, V.N., Western, D. and Tyrrell, P., 2018. Application of open source tools for biodiversity conservation and natural resource management in East Africa. *Ecological Informatics*, *47*, pp.35-44.

Nabyonga-Orem, J., Asamani, J.A., Nyirenda, T. and Abimbola, S., 2020. Article processing charges are stalling the progress of African researchers: a call for urgent reforms. *BMJ Global Health*, *5*(9), p.e003650.

Neylon, C., & Wu, S., 2009. Open Science: tools, approaches, and implications. In *Biocomputing 2009*. pp. 540–544.

OECD., 2015. “Making Open Science a Reality”, OECD Science, Technology and Industry Policy Papers, No. 25, OECD Publishing, Paris. http://dx.doi.org/10.1787/5jrs2f963zs1-en

Pardi, N., Hogan, M.J., Porter, F.W. and Weissman, D., 2018. mRNA vaccines—a new era in vaccinology. *Nature Reviews Drug Discovery*, *17*(4), pp.261-279.

Parker, T., Fraser, H., Nakagawa, S., Hampton, J.O., Warburton, B. and Sandøe, P., 2019. 747 Making conservation science more reliable with preregistration and registered reports. *Conservation Biology*, *33*(4).

Peterson, A.T., Anderson, R.P., Beger, M., Bolliger, J., Brotons, L., Burridge, C.P., Cobos, M.E., Cuervo-Robayo, A.P., Di Minin, E., Diez, J. and Elith, J., 2019. Open access solutions for biodiversity journals: Do not replace one problem with another. *Diversity and Distributions*, *25*(1), pp.5-8.

Pettorelli, N., Safi, K. and Turner, W., 2014. Satellite remote sensing, biodiversity research and conservation of the future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1643), p.20130190.

Poisson, A.C., McCullough, I.M., Cheruvelil, K.S., Elliott, K.C., Latimore, J.A. and Soranno, P.A., 2020. Quantifying the contribution of citizen science to broad-scale ecological databases. *Frontiers in Ecology and the Environment*, *18*(1), pp.19-26.

Pool, J., Fatehi, F. and Akhlaghpour, S., 2021. Infodemic, misinformation and disinformation in pandemics: scientific landscape and the road ahead for public health informatics research. In *Public Health and Informatics* (pp. 764-768). IOS Press.

Rzymski, P., Nowicki, M., Mullin, G.E., Abraham, A., Rodríguez-Román, E., Petzold, M.B., Bendau, A., Sahu, K.K., Ather, A., Naviaux, A.F. and Janne, P., 2020. Quantity does not equal quality: scientific principles cannot be sacrificed. *International Immunopharmacology*, *86*, p.106711.

Scaria, A.G. and Rangarajan, R., 2016. Fine-tuning the IP Approaches for Fostering OS: Some Insights from India. 8 WIPO Journal 109. Available at SSRN: https://ssrn.com/abstract=2844625 or http://dx.doi.org/10.2139/ssrn.2844625

Schade, S. and Tsinaraki, C., 2016. Survey report: data management in Citizen Science projects. *Publication Office of the European Union: Luxembourg*.

Simms, S.R. and Jones, S., 2017. Next-Generation Data Management Plans: Global, Machine-Actionable, FAIR. *International Journal of Digital Curation, 12,* p.36–45

Stokstad, E., 2020. UK cues up big funding increases for R&D. *Sciencemag Online*. Accessed 9th November 2020. Available at: https://www.sciencemag.org/news/2020/03/uk-cues-big-funding-increases-rd

Storksdieck, M., Shirk, J.L., Cappadonna, J.L., Domroese, M., Göbel, C., Haklay, M., Miller-Rushing, A.J., Roetman, P., Sbrocchi, C. and Vohland, K., 2016. Associations for citizen science: regional knowledge, global collaboration. *Citizen Science: Theory and Practice*, *1*(2).

Strasser, B. and Haklay, M.E., 2018. Citizen Science: Expertise, democracy, and public participation. *SSC Policy Analysis 1/2018* , pp. 1-92

Sullivan, B.L., Phillips, T., Dayer, A.A., Wood, C.L., Farnsworth, A., Iliff, M.J., Davies, I.J., Wiggins, A., Fink, D., Hochachka, W.M. and Rodewald, A.D., 2017. Using open access observational data for conservation action: A case study for birds. *Biological Conservation*, *208*, pp.5-14.

Suman, A.B. and Pierce, R., 2018. Challenges for citizen science and the EU open science agenda under the GDPR. *Eur. Data Prot. L. Rev.*, *4*, p.284.

Taylor, S.J., Anagnostou, A., Fabiyi, A., Currie, C., Monks, T., Barbera, R. and Becker, B., 2017, December. Open science: Approaches and benefits for modeling & simulation. In *2017 Winter Simulation Conference (WSC)* (pp. 535-549). IEEE.

Teixeira da Silva, J.A. and Bornemann-Cimenti, H., 2017. Why do some retracted papers continue to be cited?. *Scientometrics*, *110*(1), pp.365-370.

The Royal Society., 2019. Investing in UK R&D. *[Online]*. Accessed 9th November 2020. Available at: https://royalsociety.org/-/media/policy/projects/investing-in-uk-r-and-d/2019/investing-in-UK-r-and-d-may-2019.pdf

Toelch, U. and Ostwald, D., 2018. Digital open science—Teaching digital tools for reproducible and transparent research. *PLoS Biology*, *16*(7), p.e2006022.

Toerpe, K., 2013. The rise of citizen science. *The Futurist, 47*(4), 25.

Vicente-Sáez, R. and Martínez-Fuentes, C., 2018. Open science now: A systematic literature review for an integrated definition. *Journal of Business Research*, *88*, pp.428-436.

Williams, J., Chapman, C., Leibovici, D., Loïs, G., Matheus, A., Oggioni, A., Schade, S., See, L. and van Genuchten, P., 2018. Maximising the impact and reuse of citizen science data. UCL Press.

Williams, M., Bagwell, J. and Zozus, M.N., 2017. Data management plans: the missing perspective. *Journal of Biomedical Informatics*, *71*, pp.130-142.

Chapter 4. Determining Species-Rich Grassland Sites and Classifications to Create a Habitat Prediction Model for Mapping Species-Rich Grasslands

4.1 Introduction

4.1.1 The Importance of Species-Rich Grasslands

Globally, grassland habitat decline has been driven by expanding agricultural land-use, largely over the past two centuries (Ceballos *et al.*, 2017; Dudley *et al.,* 2020). Within Europe and, specifically, the UK, species-rich grasslands (SRGs) are amongst the most threatened habitats, covering less than 1% of the UK landscape (Pywell *et al*., 2002; NatureScot, 2021). This reduction is problematic due to the vital ecosystem services that grassland habitats provide, including as carbon storage and as habitat for declining invertebrate species (Zhao *et al*., 2020). For example, the butterfly *Aricia artaxerxes* is now listed as vulnerable from habitat loss and climate change, causing a decrease in abundance of 57% over a 40-year period (Fox *et al.,* 2023). Continued SRG and global grassland decline will result in desertification, water eutrophication from soil erosion, and further release of carbon into the atmosphere, worsening societal capability for global food security and climate change mitigation (Yan *et al.,* 2021).

4.1.2 The Need to Map Species-Rich Grasslands

There is currently an unaddressed requirement to map SRGs. Locating and mapping these habitats is crucial for their conservation and continuing ecosystem functioning. Despite the importance of SRGs, they are poorly recorded in certain areas of Europe. The European Commission implies this through the suggestions that SRGs can only be protected (from agricultural intensification, for example) by sufficiently mapping the extent of the habitat (King, 2010). Scotland, in particular, has large gaps in coverage of mapped species-rich grasslands and other habitats, with initiatives from NatureScot (including the Species on the Edge partnership with Butterfly Conservation) and the Scottish Borders Council habitat action plan identifying this as a key conservation goal (Scobie, 2018, Scottish Borders Council, n.d.).

Where areas of SRGs have been mapped, records tend to date back to the late 20th century and early 2000s (Dahlström *et al.*, 2013; Divíšek and Chytrý, 2018; Michalcová *et al.*, 2014). These areas may have undergone land use changes since records were collected and the purposes of the data collection may have differed largely from their potential use in SRG mapping. More recently, a comprehensive mapping of European habitats was undertaken using these vegetation records, which were transformed into the corresponding European Nature Information System (EUNIS) codes (Chytrý *et al.*, 2020). However, within the habitat maps, it is difficult to determine which type of grassland category is present in each plot, and overlaying the maps of broader habitat types creates confusion over which is the dominant habitat in certain areas.

Scotland has little data available on the location of SRGs, with only up to 50% of all Scottish grassland types previously mapped (Bourne, 2020). It is estimated that there are approximately 30,000 ha of

SRGs in Scotland, of which only 2.6% of undesignated locations are known (Dadds and Averis, 2014). Reasons for this paucity of data include the cost and time requirements involved in accurately mapping all SRGs across the country (Scobie, 2018). NatureScot and the Cairngorms National Park are in the process of creating a detailed habitat map of Scotland, which should contain areas of SRGs. However, it is in the early stages of development and, as such, little data is readily available (pers.comms, 2020). The process for creating this map uses previous National Vegetation Classification (NVC) surveys to accurately identify specific plant communities of habitats at a more detailed level, and then standardising this using the hierarchical classifications in the EUNIS scheme (Gov Scot, 2018). While this is the most recent and detailed habitat map of Scotland that exists, there are large gaps in coverage where habitat designations are still to occur.

Identifying areas of SRGs is central to their accurate monitoring and management, as well as for the conservation of the species dependent on the habitat. Specifically, developing an understanding of characteristics of SRG sites can inform research on why certain changes in species presence might be occurring. Because of this, methods must be explored that could provide updated maps and locations of SRGs. For example, an approach that could increase the breadth of this potential would be with remote sensing.

4.1.3 Understanding Remote Sensing for Mapping Species-Rich Grasslands

One method to achieve wide scale mapping of SRGs and other threatened habitats can be with the use of remote sensing (RS) technology, reducing labour intensity of field surveys and widening spatial reach. Advances in RS technology are increasing the potential of habitat classifications globally, specifically in relation to grasslands, which have historically been overlooked. The high intra-habitat heterogeneity and finer spatial resolution of grassland features have made it difficult for RS techniques to pick up small-scale differences between grassland classes until recently (Gholizadeh *et al.*, 2019; Wachendorf *et al.*, 2018). Determining areas of various classes of SRGs is beneficial when considering different habitat requirements of threatened species (Reddy, 2021). This has not been achieved to a great extent, with only a few studies investigating RS techniques for semi-natural grassland community mapping (Buck *et al.,* 2015; Raab *et al.,* 2018; Rapinel *et al.,* 2019; Schuster *et al.,* 2015).

Studies on grassland RS usually focus on improved agricultural grasslands, investigating crop or forage quality or quantity (Reinerman *et al.*, 2020). Only more recently has species richness of biodiverse grasslands been investigated (Imran *et al.,* 2021; Muro *et al.,* 2022; Rossi *et al.,* 2022). The global scale of these studies is not far reaching, with few studies focusing on field based remote sensing (e.g., with spectroradiometers) and satellite remote sensing over very small spatial scales – usually less than a few kilometres (Gholizadeh *et al.*, 2020 and Zhao *et al.*, 2021). As such, there is potential for RS applications to be used more widely in SRG landscapes. Specifically, creating a habitat classification model using RS data may allow predictions of unlocated areas of SRGs to be identified and feed into current mapping initiatives.

4.1.4 Building a Habitat Classification Model for Species-Rich Grasslands

The process of building a habitat classification model requires synergy across various aspects. These include a thorough understanding of environmental features of interest that can be differentiated in RS to be used as predictor variables, as well as correct ecological interpretation of habitat classes that form the classes to be determined by those predictors (Figure 4-1).

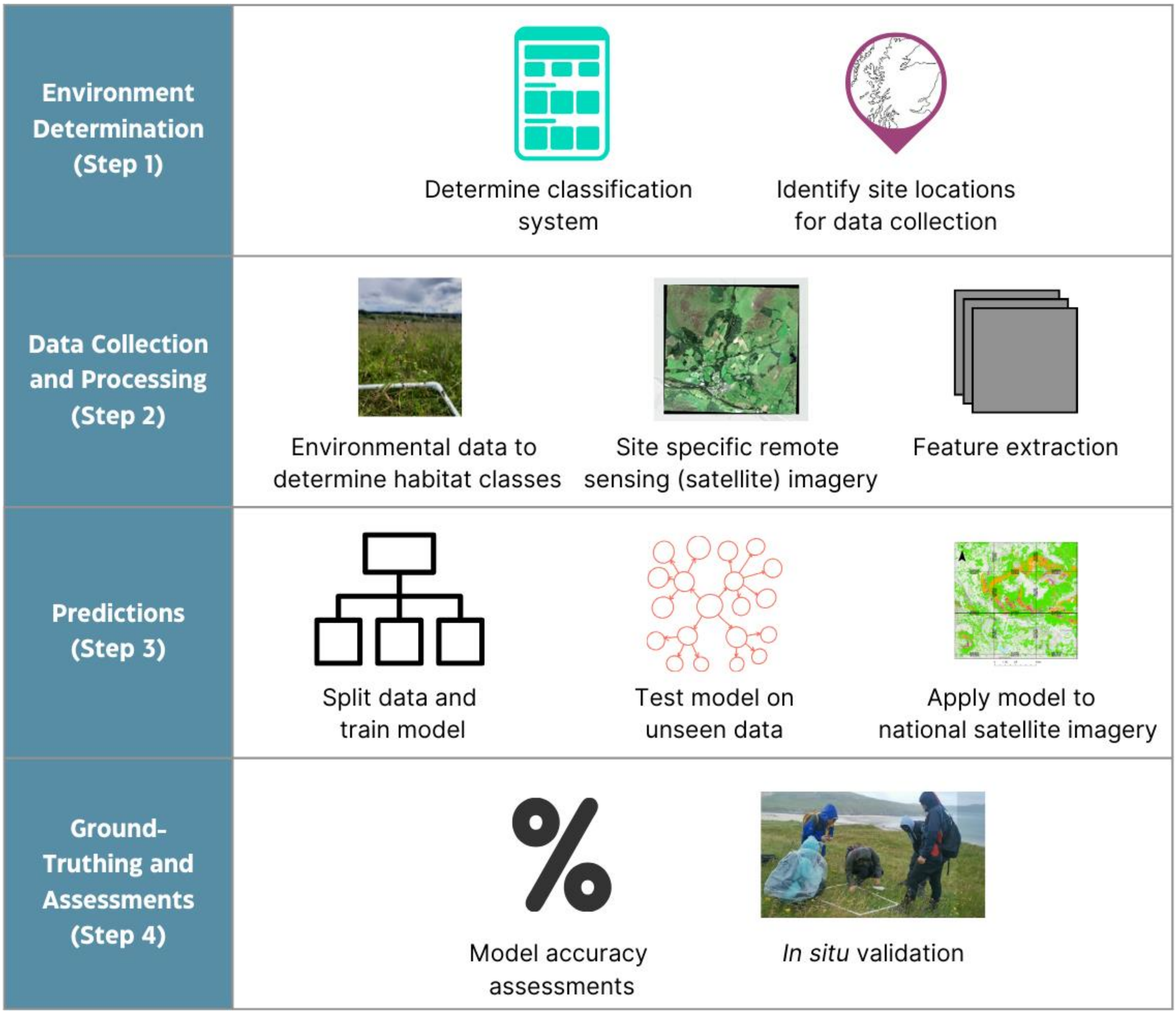


Figure 4-1. Habitat classification model workflow.

4.1.4.1 Classifying Species-Rich Grasslands (Step 1)

RS data can be fed into classification models to create prediction maps of vital habitats. However, to do so, the habitat must be correctly classified. Challenges in mapping are introduced, as the term ‘species-rich grassland’ is not one specific habitat but consists of a range of defined communities that all have a dense floral diversity. However, the definition usually removes certain grasslands on coastal or alpine areas, which would be influenced by other climatic variables (Jefferson *et al.,* 2019). In this way, it is important to highlight habitats that can be classed as SRGs to clarify the name when identifying, monitoring, and managing these grasslands.

There are a range of classification schemes that help to identify specific habitats, which can be used to characterise SRGs (JNCC, 2019a). For example, the JNCC's Phase 1 habitat classification scheme aims to identify habitats at a larger scale based on common observable characteristics (JNCC, 2010). The UK Biodiversity Action Plan utilises a similar broad habitat classification, which included associated priority habitats for protection by law (UK Biodiversity Action Plan, 2008). The NVC further categorises habitats more specifically based on their plant community makeup (Pigott *et al.,* 2000). These classification schemes are all UK based, however, they are comparable with other schemes across Europe, for example the EUNIS habitat guide, which enables standardisation of habitats across Europe (Moss, 2014).

There is often subjectivity when designating habitats, which can influence the final classification. There may also be an overlap between plant communities that can form mosaic habitats and not exclusively fit a single designation (Evans, 2006; Jefferson *et al.,* 2019). Reliable classification is essential in mapping habitats and implementing the relevant conservation methods, and the coordination of schemes at national and international scales is necessary for their wider protection (Ichter, 2014). However, there are certain species assemblages that will constrain some habitats to be specific to one country or even region. It is expected that there will always be some variance in classifications, especially considering the conservation priorities of different nations. For example, the UK requested that *Nardus* grasslands must be determined to be species-rich in the EU Habitats Directive list. This is because where the grass species *Nardus stricta* dominates the grassland (instead of a mixed grass sward including *N. stricta*), the habitat is considered improved through over-grazing, resulting in a low conservation value (Evans, 2010). Where it is possible, harmonising classifications should occur to the greatest extent.

Using these schemes, SRGs can be divided into specific categories based on characteristics, such as soil properties and climate, which influence the species diversity present on the habitats. These divisions include acid, neutral, and calcareous grasslands, which are determined by the pH of the soil (JNCC, 2010). Other broad SRG categories are marsh grasslands as outlined by the JNCC (2010), or fen/marsh/swamp as identified by the UK Biodiversity Action Plan (2008). In this grassland category there is potential for overlap across various classes in the schemes. For example, the priority habitat of 'upland flushes, fens, and swamps' as outlined by UK Biodiversity Action Plan (2008) could potentially be identified as bogs/mires (class E), swamps (class F) or marsh (class B) as highlighted by the Phase 1 habitat guide of JNCC (2010). The final classification assigned is heavily reliant on the dominant presence of certain species (e.g., mosses vs. grasses/herbs) and the seasonal variation in height of the water table (JNCC, 2010). Due to the great overlap potential with these classifications of wet grasslands, SRGs of this class will be referred to as marsh grasslands throughout this thesis, as outlined by Jackson (2000).

Each of these grassland divisions has a priority habitat associated with it, as defined in UK legislation, as well as associated NVC communities (**Error! Reference source not found.**) (Rodwell, 1992; UK B iodiversity Action Plan, 2008). These classifications and descriptions were originally published at the end of the 20th century, with few updates occurring by the next assessments. Therefore, these figures are merely estimates, and there is much more to be done in terms of monitoring and managing these habitats across the UK (JNCC, 2019b). Together, these divisions will form the basis of SRG classifications throughout this thesis, which align with the EUNIS habitat codes. These codes are used by external organisations such as Butterfly Conservation when completing records for the UK Butterfly Monitoring Scheme (UK BMS), and, as such, are important for classifying habitats for associated species (Brereton *et al.,* 2000). These classifications are required when determining site locations that will be visited to collect environmental and RS data that will feed into the model.

Table 4-1. Species-rich grassland designations across classification schemes and estimated extent and protection. (NVC codes where overlapping communities may occur). Table adapted from Jackson (2000). Descriptions from UK Biodiversity Action Plan (2008); Shellswell *et al*. (2016); Crofts and Jefferson (1999).**

Broad UK Grassland Classification	**Associated Priority Grassland Classification**	**Classification Codes**			**Description**	**Estimated Extent**	**Protection**
		BAP Phase 1	*EUNIS*	*NVC (species-rich)*			
Acid Grassland (pH <5.5)	Lowland dry acid grassland	B1	E1.7*, E1.73 E1.92	U1 – U5 *CG2 CG7 CG10 CG11 CG13*	Nutrient poor, free draining soils; both species poor (<5/4 m^2) to species rich (>25/4 m^2).	~30,000 ha in the UK	Inclusion in PAs (extent unknown); agri-environmental schemes
Neutral grassland (pH 5.5 – 7)	Upland hay meadow.	B2	E2 except 2.6	MG1 (excluding a and b) MG2 – MG5, MG8, MG11 – MG13	Dense grass growth; swards up to 80 cm in height; moderate slopes up to 400 m altitude	<2,000 ha but extent unknown. Restricted to northern UK. Extent unknown.	Majority in PAs in England, very little inclusion in Scotland (0.5%). Increase inclusion; agri-environmental schemes.
	Lowland meadow				Unimproved pastures, hay meadows, some recreational sites, roadside verges	<15,000 ha in the UK.	Inclusion in PAs (extent unknown); agri-environmental schemes; increase favourable conditions.
Calcareous grassland (pH >7)	Upland calcareous grassland	B3	E1.2, E1.26 E1.72#	CG1 – CG10 *OV37*	Lime-rich, shallow soils; Usually higher altitudes (250-300m) High species richness (>60 species/ 4 m^2)	<25,000ha in the UK. Particularly important areas for the habitat include the North Pennines and Cumbria in England and Breadalbane in Scotland	Up to 40% included in PAs; Maintain and increase condition of current extent;
	Lowland calcareous grassland				Lime-rich, shallow soils; flatter topographies; species-rich with rare species	Up to 41,000 ha in UK. None in Scotland.	Up to 70% in PAs. Heavy involvement in agri-environmental schemes.
Marsh Grassland	Purple moor grass and Rush pasture	B5	E3.42, E3.511, E3.512	M22 – M26 + M27 *U5b U6*	Acid soils; poor drainage; species-rich	<56,000 ha in the UK. Important location for this habitat across Europe.	~6.8% in PAs in the UK. Increase condition both within and outside of PAs. Inclusion in agri-environmental schemes.

4.1.4.2 Environmental and Remote Sensing Features as Predictors of Species-Rich Grasslands (Step 2)

There are various remote sensing features, that correspond to environmental factors, that will enable high level grassland classification and mapping. RS data that can capture multiple wavelengths and have higher spatial resolutions are more beneficial for grassland identification, as they can classify characteristics of vegetation with greater accuracy. For example, the yellow spectral band (575 – 585 nm) can identify how ripe a plant is whilst the red band (610 – 700 nm) can identify chlorophyll content, which is important for measuring LAI and nutrient content of plants (Ali *et al.*, 2016). The launch of Sentinel-2 (S2) in 2015 provides this opportunity, as more spectral data can be encapsulated due to its multispectral capacity, capturing data across 13 bands (ESA, 2013). The inclusion of the red-edge bands, for example, enable greater information to be received in narrower wavelengths where the chlorophyll content is detected (Clevers and Gitelson, 2012). Reflection data across more wavelengths may be the best way at differentiating the various grassland classes, with evidence demonstrating spectral diversity as a possible way to show species diversity, especially with functionally different plant species (Fassnacht *et al.,* 2022; Peng *et al.*, 2019). This is known as the Spectral Variation Hypothesis (SVH) (Palmer *et al.*, 2000). The use of RS in investigating the SVH has been further explored in chapter 5 of this thesis, to assess its potential in improving mapping attempts of SRGs after initial classification predictions. S2 also has a high spatial resolution of 10 m (in certain bands) and frequent revisitation rate (5 days), resulting in greater data capture throughout the grassland growing season (ESA, 2013). This allows detailed time series to be created, which aids grassland RS (Franke *et al.,* 2012).

By utilising multiple RS techniques, crossing different spectral and spatial resolutions, a variety of vegetation indices (VIs) may be calculated as predictor variables, which have also previously been shown to vary with species diversity (Chitale *et al.*, 2019). The Normalised Difference Vegetation Index (NDVI) is commonly used due to its ability to discriminate between areas of vegetation and other surfaces. Further indices, such as EVI (Enhanced Vegetation Index) and S2REP (Sentinel-2 Red-Edge Position) are helpful for differentiating between multiple grassland habitats. The red edge bands' sensitivity to chlorophyll may indicate healthier or more abundant grassland vegetation, whilst EVI is also commonly used due to its relation to species diversity (Bekkema and Eleveld, 2018; De Simone *et al.,* 2021; Peng *et al.,* 2019 Rapinel *et al.,* 2019; Stenzel *et al.*, 2017). Price *et al.* (2002) also demonstrated that the Greenness Vegetation Index (GVI) may be a useful addition for discriminating between grassland classes.

Selecting the most appropriate indices must be context based due to the vast range of indices available. For example, the Soil Adjusted Vegetation Index may be more beneficial in areas where bare soil is prominent, or water adjusted indices may take advantage of the short-wave infrared (SWIR) bands in wetter landscapes (Price *et al.,* 2002). Having multiple VIs for the different methods of RS is useful to ascertain how these compare in relation to accurate habitat classification (Moon *et* *al.,* 2021). Once these features have been determined, this information is aligned with each habitat class as outlined by the classification system to be used as predictor variables.

4.1.4.3 Creating, Training, and Applying a Model (Step 3)

There are various methods by which a habitat classification model could be created. Classification types include *a)* "supervised", where known features are used to calculate considerations in classifications, *b)* "unsupervised", where ground features are unknown and classes are created from similar pixel properties, and *c)* "object-based image analysis (OBIA)", where distinct objects are identified then classified (Al-Doski *et al.,* 2013). There are advantages and disadvantages to classification types, including time and experience needed for supervised classification and potential human error in this method compared to lack of input for unsupervised classifications, resulting in reduced separation and correlation to the classes of interest (Enderle and Weih, 2005).

Supervised classifications appear to be most prominent and have often provided the most accurate classifications (Boori *et al.,* 2018). There are a range of classifiers in each category used for land cover assignment, using different algorithms created from features of the image. Examples of supervised classifications include 1) Maximum Likelihood Classification (MLC) which assigns the land cover based on the most likely habitat calculated from a function of density, 2) Random Forest (RF) which creates several decision trees that are randomly split into potential classes and the class that is most frequent is assigned, and 3) Support Vector Machine (SVM) which creates a hyperplane boundary splitting the data into two categories (Adankon and Cheriet, 2009; Erdanaev *et al.,* 2022).

When the data has been processed and prepared, each habitat class will have associated values across each predictor variable. Different classifiers can then be tested to find the model with the highest prediction accuracy. Data is then split into training data, that the model is based on, and then unseen test data, where the model is first applied on. Once the final model is determined, it has the potential to be applied across further unseen areas. However, unlike with train and test data, the validation of these wider predictions will not occur computationally. Validation *in situ* is, therefore, needed.

4.1.4.4 Ground-Truthing Species-Rich Grasslands (Step 4)

To ascertain the success of RS in predicting SRGs nationally, the locations identified from classification maps must be validated or ground-truthed. This tends to be a time intensive process that is limited by surveyor availability and associated costs. These limitations can be addressed by encouraging public participation in data collection with the use of citizen science. Boyd *et al.* (2022) highlight that citizen scientists can engage in Earth observation projects by verifying outputs, ground-truthing data, adding missing information, and providing local knowledge. This potential has not been met yet but is continuously identified as crucial in reaching the sustainable development goals, as outlined in chapter 1 (Karagiannopoulou *et al.,* 2022).

Due to the issues mentioned regarding the alignment of classification schemes, applicability of ground-truthing and classifying methods across grassland mapping studies will be difficult. For example, Meng *et al.* (2022) were able to initially classify grassland classes in Inner Mongolia from field-based aerial imagery as a replacement for field-based ground measurements. This would be incredibly difficult for the SRG classes of Scotland due to the spatial resolution and similar structural profile of the various classes, requiring ground-based surveys for at least one stage of the classification and mapping process. The success of these mapping attempts has also varied, even within studies, where in Canada one grassland class was able to be classified with a high success (98.2%), whereas a mixed grassland class had a low classification success rate of 45.85% (Badreldin *et al.,* 2021). The mixed grassland was composed the two other investigated grassland classes (native and tame), and it can be assumed that this increase in variation within the class may lead to the blurring of the spectral reflectance. This would likely be problematic within the SRG classes of the UK. However, few studies have investigated the success of mapping SRGs here. As such, this requires further investigation.

4.2 Aims and Objectives

For this chapter, RQ3 was adapted to be case-study specific, resulting in the chapter's overall aim: Can a habitat classification model be created to predict species-rich grasslands in Scotland, and locate habitat for vulnerable species? To enable mapping of SRGs on a larger scale using RS techniques, individual site locations of SRGs must be identified to gather information on habitat features. This information can then be used to create training datasets for a habitat classification model to locate and classify areas of SRG through RS, and to then identify further habitat locations for priority species, such as *A. artaxerxes*. Ensuring the high quality of the training datasets is a requirement for accurate classifications of different habitat types. Not only this, but *in situ* measurements collected at sample sites are needed to support RS data to calibrate the model and improve mapping accuracy (Pause *et al.,* 2016). These measurements tend to focus on quality of the habitat to be remotely sensed, which will be appropriate for identifying grassland classes that vary in condition (Dlamini *et al.,* 2016; El-Rawy *et al.,* 2019).

As such, the specific objectives of this chapter were to:

i) Identify site locations to characterise Scotland's species-rich grasslands and habitat for *A. artaxerxes*.
ii) Define and classify species-rich grasslands across Scotland.
iii) Create a species-rich grassland habitat prediction model and assess its accuracy.
iv) Nationally apply the model to predict areas of species-rich grasslands and potential habitat for *A. artaxerxes*.

The predicted locations of SRGs resulting from this chapter were then implemented in the use of a subsequent citizen science survey, as detailed in chapter 6.

4.3 Methods

Initially, known locations of SRGs in Scotland needed to be identified. This would allow the data collection of both environmental and remote sensing data associated with Scottish SRGs. The environmental data would help to classify the various types of SRGs (e.g., calcareous, neutral, acidic etc.), based on the outlined classification system in section 4.1.4.1, which need to be differentiated for the model creation, as target species will rely on specific classes within the broader SRG habitat. For example, there is a large focus when choosing the sites on calcareous grasslands and the grassland conditions required by *A. artaxerxes*, as this was a target species for this thesis and the work with Butterfly Conservation. The environmental data could then be related to the associated RS data for each site (and its corresponding SRG class) to define the classes, which are needed as the output variables in the resulting habitat classification model. The RS data (surface reflectance values and vegetation indices, discussed below) and specific environmental variables (such as the topographical data) are then used as predictor variables in the model to ascertain each of those SRG classes.

To be able to do this, several steps firstly needed to occur to determine these site locations, which incorporated a range of grassland and climatic conditions that SRGs experience across Scotland. This was required so the variation could be captured in the RS data and hopefully increase the possibility of accurate differentiation of the SRG classes. This would further support locating specific habitat for the target species *A. artaxerxes*. The data sources that are associated with helping locate SRGs in Scotland, and determine the sites in the following section, are largely down to what was available on the known locations of SRGs, as this information was very limited and largely outdated.

After the sites were located, data had been collected, and final classifications were determined, a species-rich grassland classification model could be created. This model can be applied nationally to predict species-rich grasslands, potentially identifying calcareous grasslands within this, that is the specific habitat for *A. Artaxerxes*. The subsequent methods follow the flow of model creation outlined in Figure 4-1.

4.3.1 Selection Process to Identify Sites for Characterising Species-Rich Grasslands

Multiple sites were selected from across Scotland to account for multiple SRG classes and enable the detection of further SRGs where priority species such as *A. artaxerxes* may be found or could theoretically colonise. It was important to include SRG sites where there are areas of fluctuating populations of targeted at-risk species, as this can highlight appropriate indicators that will help identify favourable environments. There is not much data available in Scotland on sites with changing *A. artaxerxes* populations but sites could be found in Northern England (Figure 4-2). As such, possible sites in England were included to capture a greater range of SRG conditions for the most accurate habitat prediction for *A. artaxerxes*. To be able to determine these final sites, available secondary data on potential SRGs and varying climatic conditions were investigated.

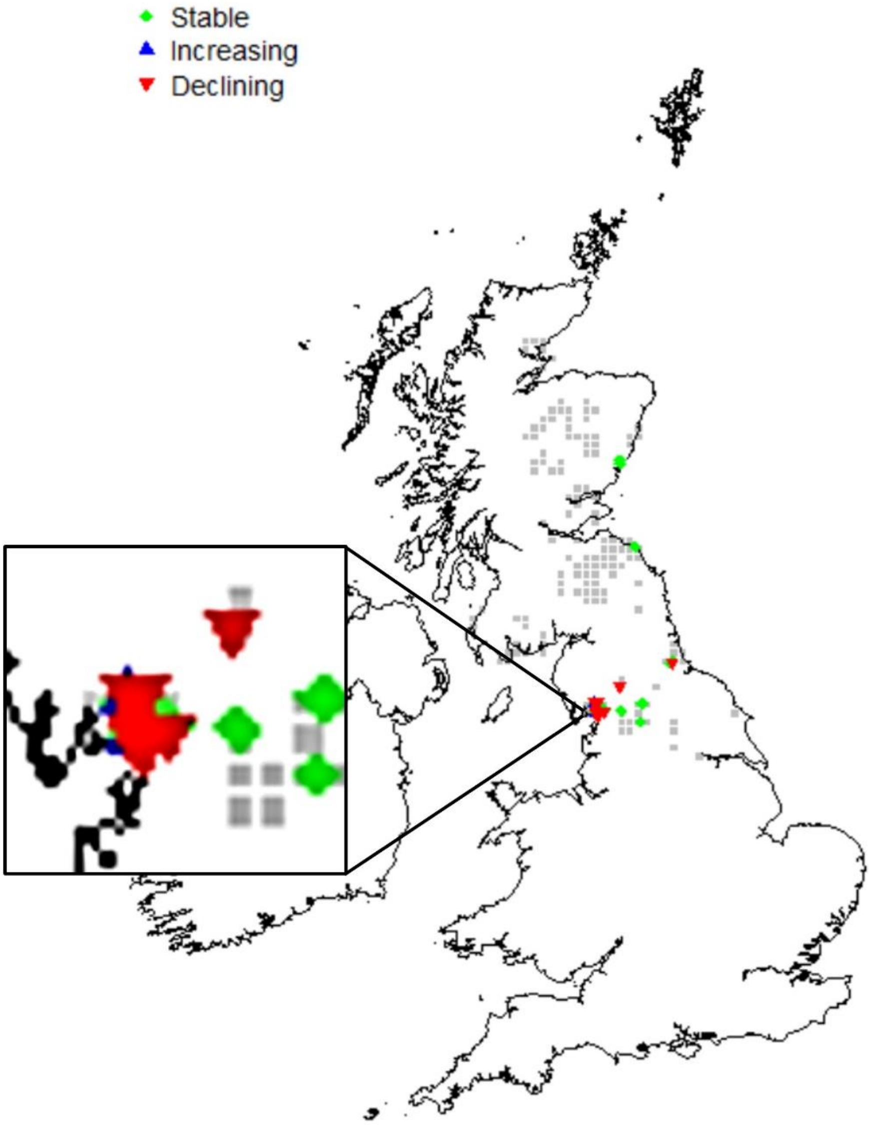


Figure 4-2. Site locations for populations of varying *Aricia artaxerxes* population trends. Green circles indicate stable populations, red arrows indicate declining populations, and blue arrows indicate increasing populations (which can be seen in the zoom insert in the North West of England). Grey cells indicate occupied sites over the past 10 years. Map from UK BMS (accessed 2020).

4.3.1.1 Identifying Potential Habitat Areas

To compare existing distribution data and select sites for analysis, secondary data containing previously mapped or collected habitat information was located from the few sources that had information on previous or potential SRG sites. ArcGIS Pro (ESRI Inc, 2020) was used for data handling. One source was the open access Habitat Map of Scotland (HabMoS). This current map shows the distribution of presently classified habitats of Scotland, made up of previous NVC surveys supplemented with other habitat surveys, and categorised into corresponding EUNIS codes. The map was downloaded as a shapefile via NatureScot (previously Scottish Natural Heritage) (Appendix C-1 for citation). The data was sorted and extracted via its attribute table into different SRG types using the associated EUNIS codes that correlate to these habitats, as outlined in **Error! Reference source not f**

ound.. These feature layers were merged to create one continuous layer of potential areas of SRGs (NS_Grasslands) that may be presently mapped by NatureScot.

Further data was acquired on land cover use in the UK, that was accessed from Edina Digimap (see Appendix C-1 for citation). This land cover map was created by the UK Centre for Ecology and Hydrology (UK CEH) to link historical records and satellite imagery to the UK BAP broad habitat scheme. In this dataset, semi-natural grasslands were classed based on the combination of habitats that were outlined as acid, neutral, calcareous, and fens/marsh/swamps (Morton *et al.,* 2020). This data was converted from raster to vector format to be able to extrapolate the specific habitat classes identified above. Features were selected based on the Morton *et al.'s* (2020) guidance, outlined to identify SRG habitats, and extracted (Table 4-2). These layers were merged to create one continuous layer of potential areas of SRG for Scotland (CEH_Grasslands) as was done with the NS_Grasslands above.

The NS_Grasslands and CEH_Grasslands were then visually compared to show similarities in the distribution and extent of potential areas of SRG located via different methods. These two map layers were also merged to create one continuous extent map of potential SRG presence (Potential_SRG), to be incorporated with further data below for site selection.

Table 4-2. Grassland classifications with associated feature codes (LC Identifiers). Used by the UK Centre for Ecology and Hydrology to seperate UK BAP broad habitats. Table adapted from Morton *et al.* (2020).

UK CEH Aggregate Class (AC)	UK BAP Broad Habitat	LC Identifier (feature code for extraction)
Semi-natural grassland (not classified by UK CEH as species-rich as includes species-poor areas)	Neutral grassland	5
	Acid grassland	6
	Calcareous grassland	7
	Fen, marsh, swamp	8

Other available data on SRGs in Scotland was the database of lowland grassland sites produced by NatureScot. This database included specific sites of certain NVC communities that are classified as SRGs. A data request was set up with NatureScot to access this information as a CSV file (Appendix C-1 for citation). To display the data, Ordnance Survey (OS) grid references for the sites were converted to latitude and longitude and mapped in ArcGIS Pro. A new layer of high value lowland grassland sites (HVLG) was created. This layer was superimposed with the other two data layers to further identify site locations. These three layers were used to show the difference in possible SRG capture from a variety of sources (Table 4-3).

Table 4-3. Difference in extent (in hectares) of potential species-rich grasslands between various map sources - UK CEH and NatureScot. Appendix C-1 for citations.

ArcGIS Pro Layer	Layer Source	Date Obtained	Approximate Habitat Area (ha) After Extraction	Projection	Layer Type
CEH_Grasslands	UK CEH	07/12/2020	1,216,444	British National Grid	Raster converted to Vector
NS_Grasslands	NatureScot	07/12/2020	120,829.4	WGS 1984	Vector
HVLG	NatureScot	13/01/2021	8255.9	WGS 1984	Point

4.3.1.2 Climate Data

Available climate data was sought to compare climatic conditions across sites and ultimately include sites of these varying conditions and see how this may influence SRG state and, as such, *A. artaxerxes* presence. Climate data on the 30-year mean (1981-2010) annual precipitation, temperature, and wind for the UK at 1 km resolution were obtained via the Met Office (Appendix C-1 for citation). Datasets were converted from a netCDF file to a float raster. These layers were then converted to an integer raster and finally to polygons. Values of zero were excluded from the datasets, as these indicated grids that were largely over sea areas and had no information. Data to be considered for site inclusions were temperatures, rainfall, and wind combinations that had relative low values (5 – 7 ° C, <1500 mm, and <5 knots) and high values (8 – 10 ° C, >1500 mm, and >5 knots). These values were chosen to ensure a wide range of varying site conditions that may influence *A. artaxerxes* presence.

4.3.1.3 Species Data

Distribution data on *H. nummularium* was acquired to locate sites where *A. artaxerxes* larval food is found. This data was also used to identify sites where the food plant exists but *A. artaxerxes* has not been observed. For example, in the west of Scotland there is an area of *H. nummularium* presence where the distribution of *A. artaxerxes* does not extend to (Figure 2-1). *A. artaxerxes* distribution data was also required to relate both occurrence datasets, as well as for analysis purposes of climatic requirements. Distribution datasets for both species were downloaded from NBN Atlas (Appendix A-1 for *A. artaxerxes* and Appendix A-2 for *H. nummularium*). These were input into ArcGIS Pro as point data.

4.3.2 Final Site Selection

After areas of possible SRGs were mapped, site selections from within the identified area were chosen, mainly in alignment with known Butterfly Conservation transect locations. A representative sample selection was used to ensure sites were not chosen where SRG is not present, and to include sites of differing SRG conditions (such as varying climates and grassland soils), to assess their influence on priority species distributions. Due to the large number of transect locations in Scotland, certain sites were excluded to narrow down the selection that would be most relevant to locating areas of SRG for conserving *A. artaxerxes*. Butterfly surveying sites must have been found within potential SRGs and below 350 m altitude. Sites also must have been found in areas with average annual temperatures above 4° C, below 2500 mm of rainfall, and with wind speeds of less than 10 knots, as *A. artaxerxes* has not been recorded in Scotland outside of these, based on the analysis in this thesis. Climatic requirements of *A. artaxerxes* appear not to be described in literature in the wider European context either.

The distribution of *A. artaxerxes* across temperatures (° C) in Scotland is summarised (Table 4-4) and displayed (Figure 4-3).

Table 4-4. Distribution of *Aricia artaxerxes* summarised by 30 year mean annual temperatures (° C) across Scotland.

Temperature (° C)	1	2	3	4	5	6	7	8	9	10
Count	0	0	1	16	50	322	1249	947	950	10

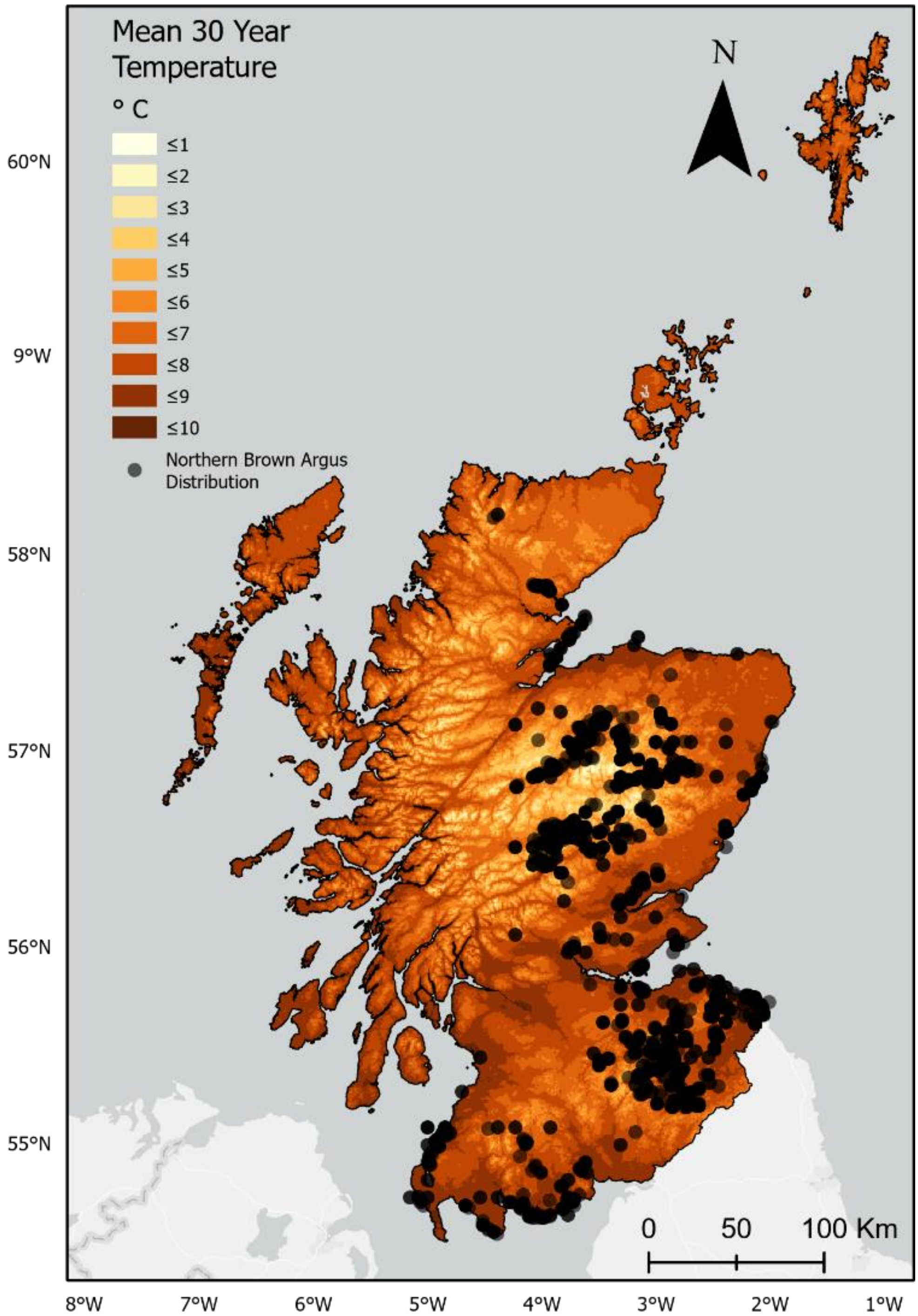


Figure 4-3. Distribution of *Aricia artaxerxes* across annual 30 year (1980-2010) mean temperature in Scotland. Temperatures were displayed after conversion from raster grid to polygon grid. Climate data accessed via the MetOffice (Appendix C-1 for citation). Aricia artaxerxes distribution data was accessed from NBN Atlas (Appendix A-1 citation).

The distribution of *A. artaxerxes* across precipitation (mm) in Scotland is summarised (Table 4-5) and displayed (Figure 4-4).

Table 4-5. Distribution of *Aricia artaxerxes* summarised by 30 year mean annual precipitation (mm) across Scotland.

Precipitation (mm)	≤892	≤1124	≤1362	≤1605	≤1866	≤2159	≤2475	≤2829	≤3248	≤4291
Count	1534	1279	475	185	54	15	3	0	0	0

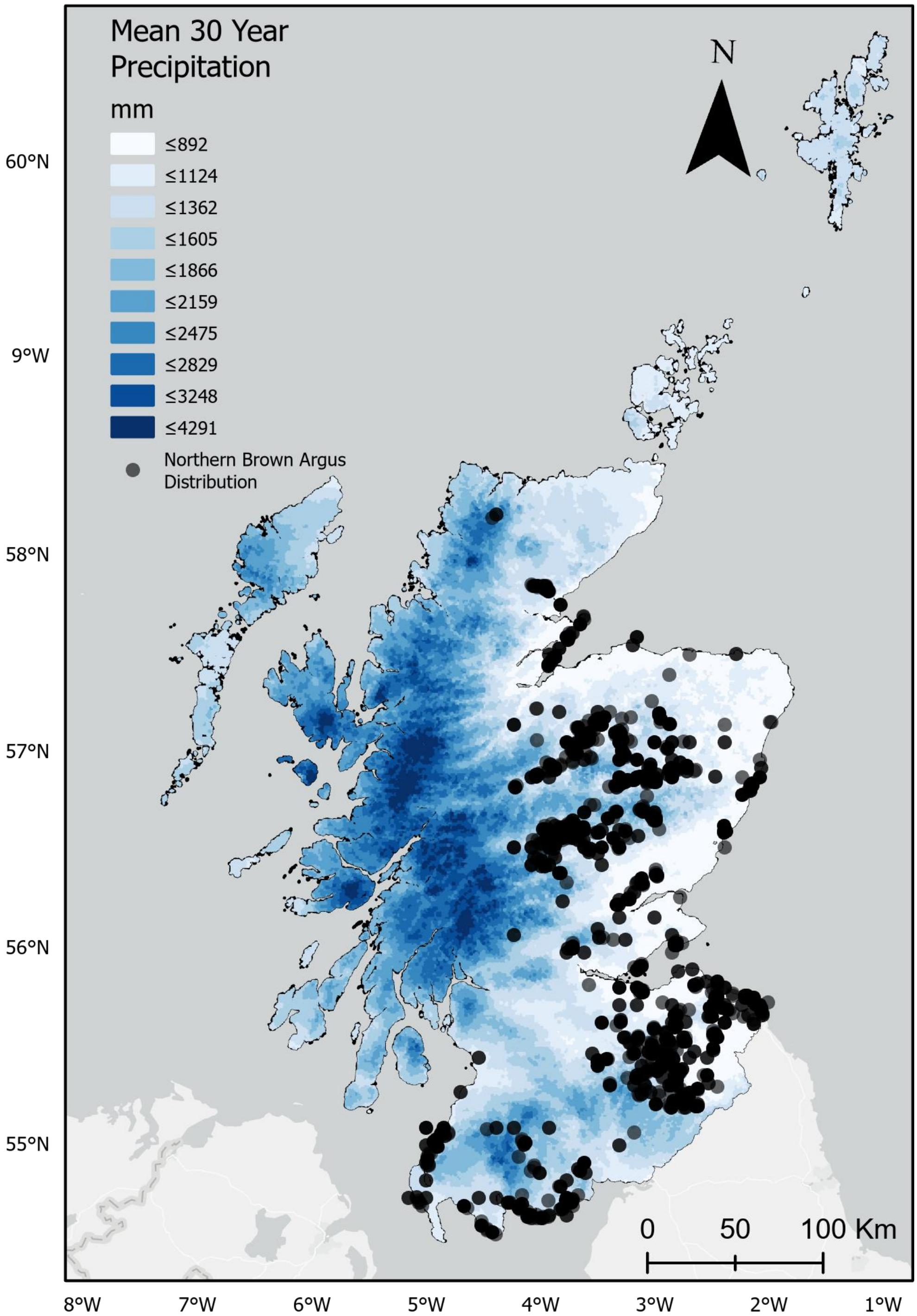


Figure 4-4. Distribution of *Aricia artaxerxes* across annual 30 year (1980-2010) mean precipitation (mm) in Scotland. Rainfall was displayed after conversion from raster grid to polygon grid. Climate data accessed via the MetOffice (Appendix C-1 for citation). *Aricia artaxerxes* distribution data was accessed from NBN Atlas (Appendix A-1 for citation).

The distribution of *A. artaxerxes* across wind speed (knots) in Scotland is summarised (Table 4-6) and displayed (Figure 4-5).

Table 4-6. Distribution of *Aricia artaxerxes* summarised by 30 year mean annual wind speed (knots) across Scotland.

Wind Speed (knots)	1	2	3	4	5	6	7	8	9	10	11	12	13
Count	3	10	402	1190	1213	656	50	22	0	4	0	0	0

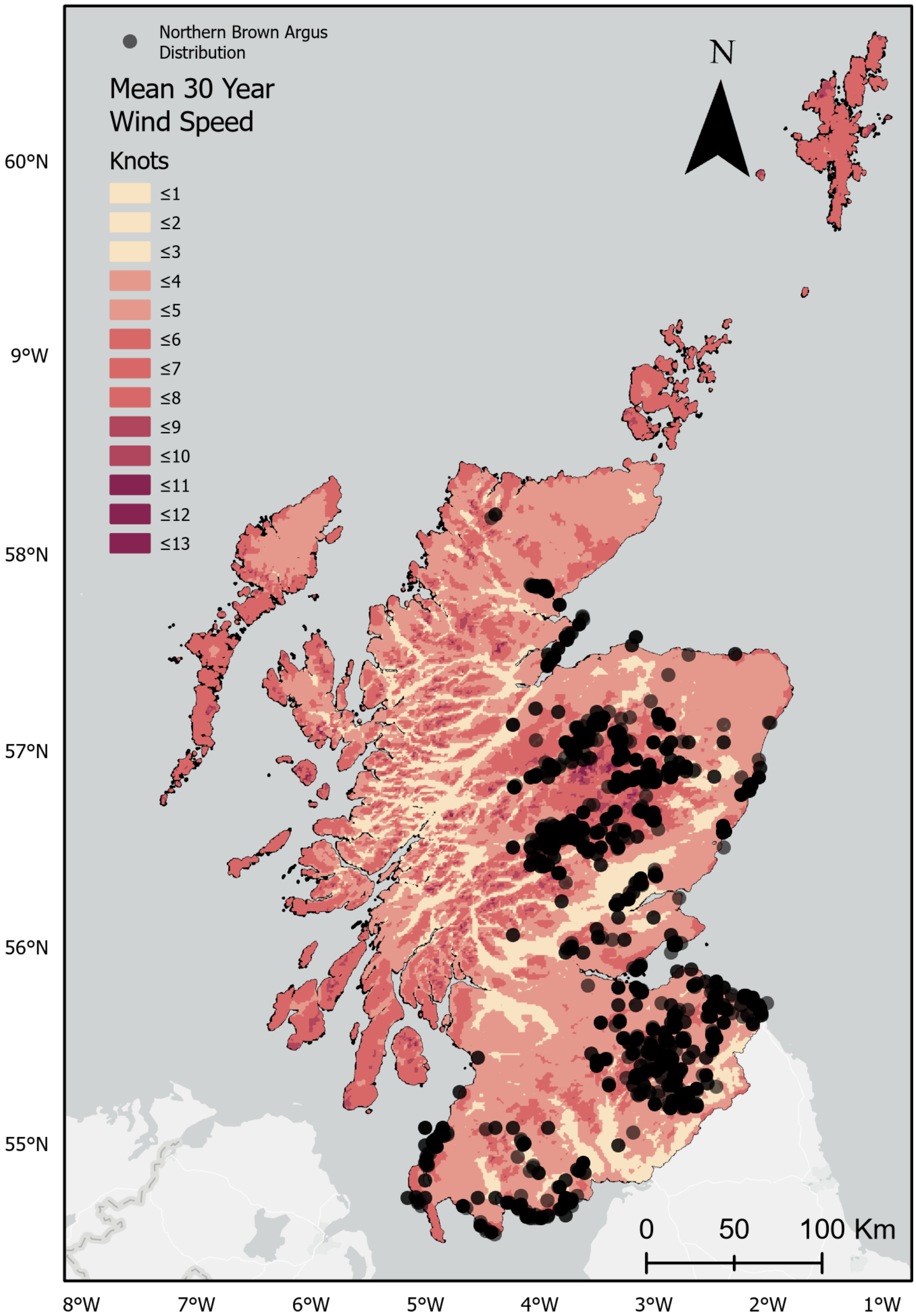


Figure 4-5. Distribution of *Aricia artaxerxes* across annual 30 year (1980-2010) mean wind speed (knots) in Scotland. Wind speeds were displayed after conversion from raster grid to polygon grid. Climate data accessed via the MetOffice (Appendix C-1 for citation). *Aricia artaxerxes* distribution data was accessed from NBN Atlas (Appendix A-1 for citation).

Site locations were obtained via transect locations from the UK Butterfly Monitoring Scheme (UK BMS) and Butterfly Conservation as CSV files (Appendix C-1 for citation). This information was used to distinguish the transects that fall within the areas of the collated Potential_SRG layer. As this data contains specific site names and exact locations, these were used for the final site selections. Multiple datasets were accessed for butterfly transect locations for an initial broader range: Site locations for 2019 (UKBMS_2019), the most recent transect information, and transects with known habitats were obtained (UKBMS_Habitats) (Table 4-7). Both layers were needed as they encapsulate different areas across Scotland. OS grid references were converted to longitude and latitude coordinates and the XY data was displayed as point locations. Initially, sites were excluded from UKBMS_2019 if they had not been surveyed within the last 5 years to incorporate the most recent data. Sites were then selected if they were included in the Potential_SRG layer.

As the site location data is collected as grid references there was the possibility that the coordinates do not align exactly due to GPS and map calibration discrepancies. To account for this, a buffer zone of 100 m was created to capture sites that may be omitted. This size was chosen as the OS grids are 1 km by 1 km. Sites outside of the buffer zone were excluded. As this layer may incorporate grassland that is not classified as SRG, the UKBMS_Habitats layer was further used to locate sites by filtering the recorded primary habitat to SRG types.

Site selection was further supported with calibration of the site locations of high nature-value from the lowland grassland database (HVLG) where the transect locations did not include the greatest range of site conditions. Overlapping sites where *H. nummularium* were present were also considered. Data on specific *A. artaxerxes* site survey locations was acquired from Butterfly Conservation's Northern Brown Argus in the Scottish Borders project. This dataset included sites where *H. nummularium* has been recorded but *A. artaxerxes* has not. This data can help determine what other factors influence the butterfly's presence and potentially the condition of SRGs. This dataset also helped locate further sites along the Scottish borders where *A. artaxerxes* has a stronghold, as well as SRG sites with a range of pressures and management. The differences in their extent and areas of coverage can be visualised in Figure 4-6.

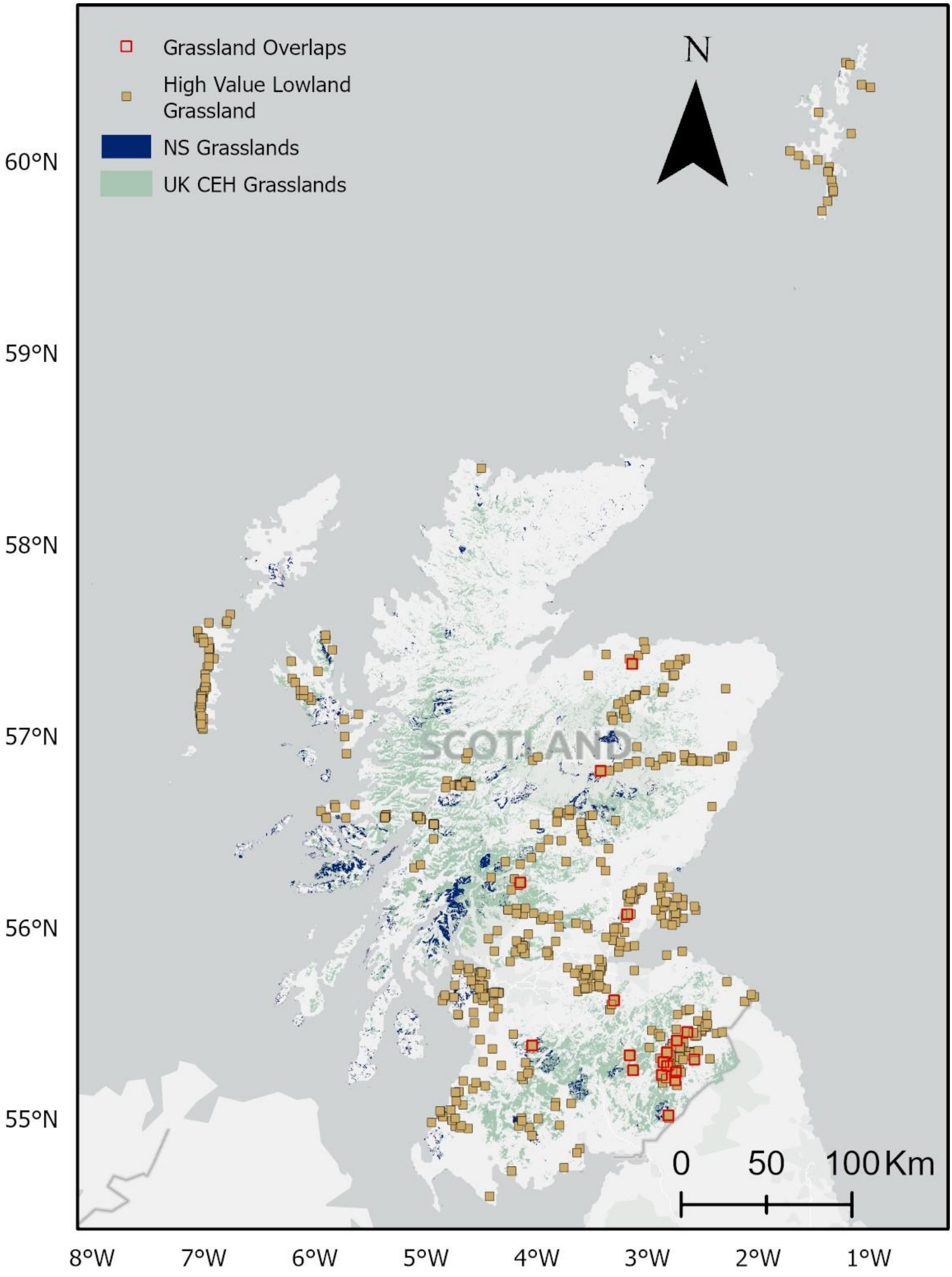


Figure 4-6. Difference in extent of areas where potential species-rich grasslands might be in Scotland. Legend: CEH Grasslands layer was extracted from the UK Centre for Ecology and Hydrology Landcover 2019 map (light green), NS Grasslands layer was extracted from the NatureScot HabMoS map (blue), HVLG points were extracted from the Lowland Grassland database from NatureScot (brown) (Appendix C-1 for citation). The area of overlap between the three data sources is outlined in red.

Site selections were also considered for circumstances not included in the previous criteria. For example, site specific information detailing *A. artaxerxes* population trends are hard to come by. Surveying that has been undertaken for *A. artaxerxes* does not provide information on whether site populations are increasing, decreasing, or stable in Scotland. This is because some sites have not had long-term monitoring (at least 5 years) to establish this data. There are few site locations in Northern England that have recorded decreasing or increasing *A. artaxerxes* populations that should be considered in the selection process. Exact site names and locations were accessed from the UK BMS website.

Table 4-7. Datasets used for site selections. See Appendix C-1 for citation.

ArcGIS Pro Layers	Data Set	Layer Source	Date Generated/Last Updated	Date Obtained	Projection
UKBMS_Habitats	UK BMS transect locations with habitat type	Butterfly Conservation	2020	02/02/2021	WGS 1984
UKBMS_2019	UK BMS most recent (2019) transect locations	UK BMS	2020	08/01/2021	British National Grid
NBA_Sites	Northern Brown Argus site locations (NBA Scottish Borders Project)	Butterfly Conservation	2021	12/01/2021	WGS 1984
HVLG	Lowland Grassland Database	NatureScot	1980s-2000s	13/01/2021	WGS 1984

4.3.3 Field Campaigns for Environmental Measurements to Characterise Species-Rich Grasslands

Three field campaigns were conducted annually at the chosen sites, timed at the beginning (early May), middle (June/July), and end (late August) of the grass growing season between 2021 – 2022. These dates were selected to detect any temporal phenological variability in grassland species. The initial field campaign was delayed due to COVID-19 restrictions and a late start to the growing season in response to cold weather. This resulted in the first campaign being combined with a reconnaissance of the sites for visual assessment to ensure they were areas of SRGs using the UK BAP Phase 1 habitat survey method. A random sampling approach was adopted to collect *in situ* measurements across the selected sites.

The transect methodology was used to take vegetation measurements for standardisation with Butterfly Conservation and butterfly recording techniques. Due to the small sizes of some of the sites, a 250 m transect (within the suggested minimum section split for butterfly transects) was used with 50 x 50 cm plot measurements taken at 50 m intervals along the transect. Intervals were chosen based on the 5-section split for butterfly transects to be able to both represent any vegetation or soils changes as well as considering feasibility due to time requirements (Schwieder *et al.*, 2020; UKBMS, 2019). Plot size was chosen as a standard in vegetation sampling along transects to provide enough detail (Boegh *et al.*, 2013; Yang, 2012). Transects were laid out in a zigzag formation to get a representative sample across the habitat and imitate butterfly flight paths (Kral, 2018; Nevis Landscape Partnership, 2015; Pellet *et al.*, 2012). This shape also emulates 'W' shaped transects used for recording vegetation measurements (Byrne *et al.*, 2018). The starting point for transects was chosen by throwing the quadrat at random.

GPS locations were taken at the start of each transect and at each centre of the subsequent quadrats using a handheld Garmin GPSMAP 62st, accurate to 3 m. This accuracy is high enough to correspond to the chosen satellite's spatial resolutions (S2's spatial resolution of 10 m). Visual measurements that were collected at sites of interest for the quadrat measurements were species presence and richness using grass identification guides. Photos of each plot were taken as well for reference. Structural measurements that were recorded include biomass, measured by taking cuttings of all AGB to grazing level within the transect plots using handheld shears. Each sample was weighed at the end of each survey providing the fresh weight (bag included). Sward height was measured randomly five times within a quadrat using a tape measure and the tallest part of the vegetation that touches the tape was recorded (Ali *et al.*, 2016; Psomas *et al.*, 2011; Shen *et al.*, 2008). The five measurements were then averaged per quadrat.

Soil measurements included pH, moisture, texture, and bulk density. Standard soil cores were taken from the top 10 cm of soil where plant species influence the soil the greatest, and organic carbon levels are found with low levels of inorganic carbon. One soil sample was taken at random in each plot along transects and the fresh weight was recorded at the end of each survey session (including core weight). All samples were taken back to the lab for further analysis. A Stevens HydraGo soil probe was used in field, taking three randomly located measurements within each quadrat to measure moisture content and averaged from the three readings. Small separate samples of soil were also collected in field to measure soil texture. Texture analysis was conducted by dampening the soil and hand manipulating for analysis using the protocol outlined by Natural England (2008). Soil bulk density was also calculated as the dry weight of the soil core divided by the volume of the soil core.

As the project works closely with the Butterfly Conservation, extra added value data was collected where necessary and applicable. This included, but was not limited to, crude habitat assessments and condition and information on food species (for example is *H. nummularium* present or not). Finally, any butterfly observations were recorded within a transect belt of 5 m either side (as is standard for the butterfly transects).

4.3.4 Laboratory Processing of Environmental Measurements

The grass biomass samples were brought back to the lab and dried at 70° C for 24 hours until a constant weight was reached and weighed again to give dry AGB, calculated per quadrat. Temperatures did not exceed 80° C to avoid combustion of the plant matter (Psomas *et al.,* 2011).

Soil cores were dried in the lab at 105° C for 24 hours until a constant weight was reached. This was to allow moisture to be calculated by boiling the water, but not burning the soil sample and combusting any organic matter within the soils. Dried soil cores were ground using a mortar and pestle and sieved with a 1.75 mm sieve to remove any plant matter, rocks, or debris for further analysis.

From the dried soil cores, 2 g of soil were removed to be mixed with 10 ml of distilled water and one drop of calcium chloride. This is the most accurate method of measuring pH and is reflective of what plants experience in the soil. The samples were magnetically mixed, and a pH probe (MettlerToledo SevenCompact pH meter) was used to measure the pH.

Organic soil carbon concentrations were measured from burning 1 g of the dried soil at 450 ° C for four hours. This temperature allows the volatile organic content to be burnt from the inorganic matter. The burnt samples were weighed, and organic content was calculated as a percentage difference (Manning *et al.*, 2015).

4.3.5 Site Ancillary Data to Improve Species-Rich Grassland Classifications

Other site information that was collected included soil texture, elevation, slope, and altitude (Scobie, 2018). Elevation data was collected from the GPS waypoints for each quadrat. Ordnance Survey 5 m Digital Terrain Models (DTM) were downloaded for each site via Edina Digimap. Slope and aspect of each quadrat location were calculated in ArcGIS Pro. Extra class data (e.g., woodlands, water, buildings, improved grasslands) were acquired through random point extraction to improve the classification model and further discriminate classes. This was particularly important to help discern the SRG classes from improved grasslands, which may have similar spectral values when improved grasslands have a high chlorophyll content before being cut/mown/heavily grazed.

4.3.6 Habitat Classification Schema to Define Species-Rich Grasslands

Species occurrence data was used to determine the final habitat classification schema for the remote sensing habitat prediction model. Visual classifications were determined using field guides, such as those by the Field Studies Council, which are accessible to all. The classification naming convention was related to the Phase 1 broad habitat classifications with inputs from the EUNIS and NVC classification schemas. Due to time intensity, complete species frequency counts were not conducted and, therefore, strict NVC classification could not be performed. NVC communities were researched post-fieldwork and identified as potential possibilities to further confirm the broad habitat types.

Environmental feature importance was used to assess how different classification schemas affect the final classification, and specifically how influential pH was on determining SRG class. Indicator species were also noted for different SRG classes on sites and suggest the possible NVC communities. Extra value data was also included to eliminate other habitat classifications e.g., woodland, as well as provide information on priority species, such as common rock-rose. Locations of these known habitats/features were determined from 1 m spatial resolution aerial imagery in ArcGIS Pro. Point locations were created for these new background classes: artificial surface, bare soil, improved grassland, water, and woodland for spectral reflectance extraction after satellite imagery acquisition.

4.3.7 Satellite Imagery Acquisition for Model Creation

S2 atmospherically corrected surface reflectance images (Level-2A) were acquired through the Google Earth Engine (GEE) platform. Images were filtered by acquisition date corresponding to the survey month for each campaign. Images were sorted by cloud pixel percentage to determine the best images where the site locations were not obscured by cloud cover. Final images for each site and campaign were chosen based on the acquisition date closest to the survey date in relation to optimal cloud cover. Full avoidance of cloud cover was not always possible, and, in some cases, clear site views required more distant acquisition dates (Appendix C-2). Specific S2 bands (B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12) that are associated with vegetation characteristics and typically used in classification analysis were selected for use in image classification and resampled to 10 m resolution. Subsets of images were created to focus more closely on the areas surrounding the site locations. Images were exported for analysis outside of GEE. The Raster package in R (v3.6.3) (Hijmans, 2023) was used to determine raster spectral pixel values for the quadrat locations in each site per campaign. Spectral reflectance data was gathered for the red, green, blue, near infrared (NIR), red-edge, and SWIR wavelengths. These satellite images were also used to extract the spectral reflectance values of the background classes, mentioned above, to enhance the classification process.

4.3.8 Vegetation Indices Calculations for Model Creation

Vegetation Indices (VIs) were calculated from the extracted pixel values of the remotely sensed images. The VIs chosen were based on previous studies using specific VIs that are related to vegetation characteristics (Table 4-8). Having the VIs for the different methods of remote sensing is useful to ascertain how these compare in relation to accurate habitat classification (Moon *et al.*, 2021).

Table 4-8. Vegetation Index Calculations.

Vegetation Index	Calculation	Reference
NDVI	(NIR-Red)/(NIR+Red)	Imran *et al.*, 2020
S2REP	S2REP = 705 + 35*((Red+ VNIR3)/2-VNIR)/(VNIR2-VNIR) using Sentinel-2 Band 4 (Red), Band 5 (VNIR), Band 6 (VNIR2) and Band 7 (VNIR3).	Li *et al.*, 2021
GVI	(NIR−GREEN)/(NIR + GREEN)	Peciña et al., 2021

4.3.9 Creating a Species-Rich Grassland Prediction Model and Accuracy Assessment

Time series spectral reflectance values from S2 across bands 2,3,4,5,6,7,8, 8A, 11, and 12 for each class were input into a csv file. VIs were calculated for each class, and ancillary data was extracted to add to the csv file for model creation in R. The data was randomly split into train (70%) and test (30%) sets to train the model on the predictor variables, used to test and apply to wider satellite images across Scotland. Two widely used classifiers, Random Forest (RF) and Support Vector Machine (SVM), were tested for grassland mapping predictions. RF was chosen as a very common classifier that has a range of applications, whilst being easy to use and understand (Belgiu and Drăguţ, 2016). SVM was investigated as is often reported as having the highest classification accuracies (Yousefi *et al.*, 2022). Accuracy assessments were performed using confusion matrices to assess which classification allows the greatest accuracy of grassland mapping. Once the model was refined, it was applied to the wider S2 image extents to predict areas of different grassland classes for citizen scientists to validate the predictions.

4.4 Results

4.4.1 Final Site Selection of Species-Rich Grasslands across Scotland

A range of sites were chosen to include multiple conditions of SRGs (Table 4-9/Figure 4-7). Due to time and access constraints, only 16 of the finalised 26 sites were able to be visited. These 16 were specifically chosen to maintain geographic range, include varying climatic conditions, and incorporate known areas of *A. artaxerxes* occurrence versus areas where *H. nummularium* is found but *A. artaxerxes* is not.

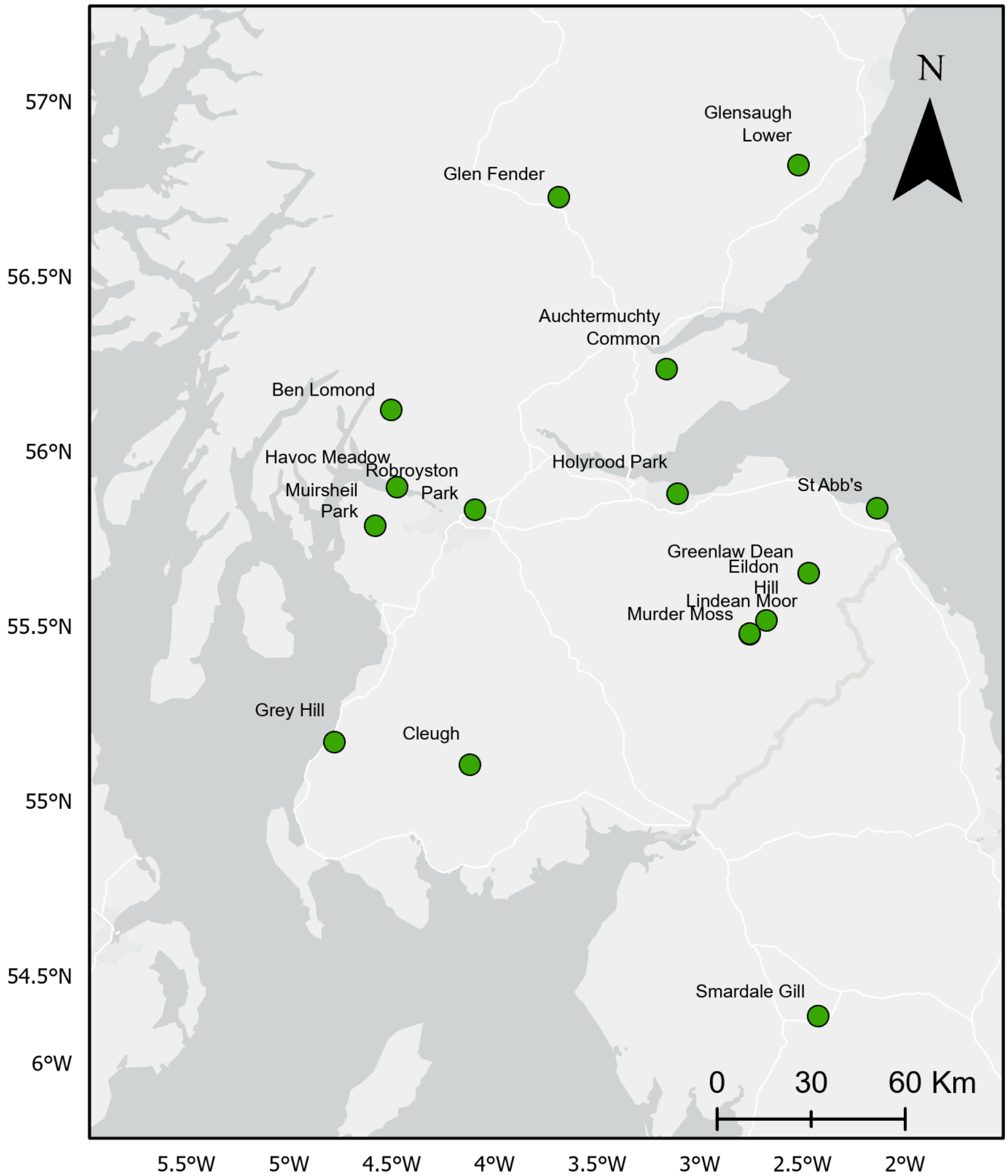


Figure 4-7. Distribution of final site locations across Scotland and North-western England for *in situ* measurements, surveyed in 2021. NB: Due to the proximity of the sites Lindean Moor and Murder Moss, they are represented as a single circle on the map, resulting in 15 location points but 16 site names.

Table 4-9. Potential site locations for in situ measurements across Scotland and North Western England. Conditions highlighted for their selection reason. Final surveyed site locations visited in 2021 highlighted in bold. Key: Grassland type - CG = calcareous, MG = neutral, U = acid. Climate - H = high, L = low, T = temperature, P = precipitation, W = Wind. Other Conditions - NBA = Northern Brown Argus, RR = Common Rock-rose.

Site Number	Site Name	Grid Ref	Grassland Type	Climate	Other Conditions
1	Arnside Knott	SD454786	CG	LP, LW, HT	Increasing NBA.
2	Warton Crag	SD494730	CG	LP, LW, HT	Decreasing NBA
3	Foulden Burn	NT921549	CG/MG	HT, LP, LW	RR present but no NBA
4	**Holyrood Park**	**NT275733**	**CG**	**LP, HT, HW**	**High RR presence intersect**
5	Megget Reservoir	NT1925	Unknown	LT, HP, HW	NBA stronghold
6	**Eildon Hill North**	**NT558328**	**U**	**HT, LP, LW**	**NBA stronghold**
7	**Ben Lomond - Ptarmigan path**	**NN362001**	**U**	**LT, HP, LW**	**NBA distribution does not reach**
8	**Glen Fender**	**NN895678**	**MG**	**LT, LP, LW**	**NBA, RR presence with colder conditions**
9	**St Abbs Head**	**NT913687**	**CG**	**HT, LP, HW**	**NBA stronghold**
10	**Havoc Meadow**	**NS381753**	**MG**	**HT, LP, LW**	**NBA distribution does not reach**
11	**Robroyston Park**	**NS629681**	**M**	**HT, LW, LR**	**NBA distribution does not reach**
12	**Lindean Moor/Whitlaw Mosses**	**NT505285**	**M/MG**	**HT, LP, LW**	**High value grass intersects**
13	**Muirshiel Country Park**	**NS311631**	**Unknown**	**LT, HP, HW**	**NBA distribution does not reach**
14	Killean Lismore	NM849417	CG	HT, HP, LW	High RR presence no NBA occurrence
15	Acreknowe	NT506112	Unknown	HT, LP, LW	High Value grass intersect; NBA stronghold site; grazed management, threat of overgrazing, RR presence but no NBA recorded
16	Earshaig	NT048026	Unknown	HT, HP, LW	Gap in NBA distribution across borders
17	**Cleugh**	**NX613867**	**MG/CG/U/M**	**HT, HP, LW**	**RR low NBA occurrence border gap**
18	Barscaigh Hill	NX863576	M/MG	HT, LW, HP	Southwest coastal area with NBA presence
19	Fallin Bing & Wester Moss	NS836910	MG	HT, LW, LP	Area of varying climatic conditions
20	**Glensaugh Lower**	**NO660780**	**U**	**LP, HT, LW**	**Gap in NBA distribution**
21	**Auchtermuchty Common**	**NO240131**	**MG**	**HT, LP, LW**	**Low NBA occurrence**
22	Glen Strathfarrer	NH368398	U	HP, HT, LW	NBA or RR don’t extend
23	Lauder Burn	NT519462	Unknown	HT, LP, LW	NBA site unmanaged no threat, grassland type unknown
24	Wurlus Burn	NT34192837	Unknown	LP, LT, LW	No threats but grazed management
25	Lealt	NG507607	MG	LW, HT, HP	High value grassland but no NBA or RR extending
26	**Grey Hill Grasslands**	**NX181941**	**CG/U**	**HT, HW, LR**	**High value grass, RR and NBA presence in borders west coast**

4.4.2 Site Conditions and Environmental Characteristics of Species-Rich Grasslands

The initial environmental data showed high trait variability both within and between sites, fluctuating largely across the survey season (Figure 4-8). Species richness ranged from 2 to 27 species found in a quadrat. Average sward height of a quadrat ranged from 5.2 cm to 115.24 cm. AGB ranged from 4.4 g/m^2 to 349.8 g/m^2 of a quadrat. Grass fresh weight ranged from 4 g to 441.9 g. Average soil moisture of a quadrat ranged from 0% to 82.37%. Average soil organic carbon content of a quadrat ranged from 4.4% to 89.41%. Average soil pH of a quadrat ranged from 3.15 to 7.03. Finally, average soil bulk density of a quadrat ranged from 0.06 g/cm^3 to 1.4 g/cm^3.

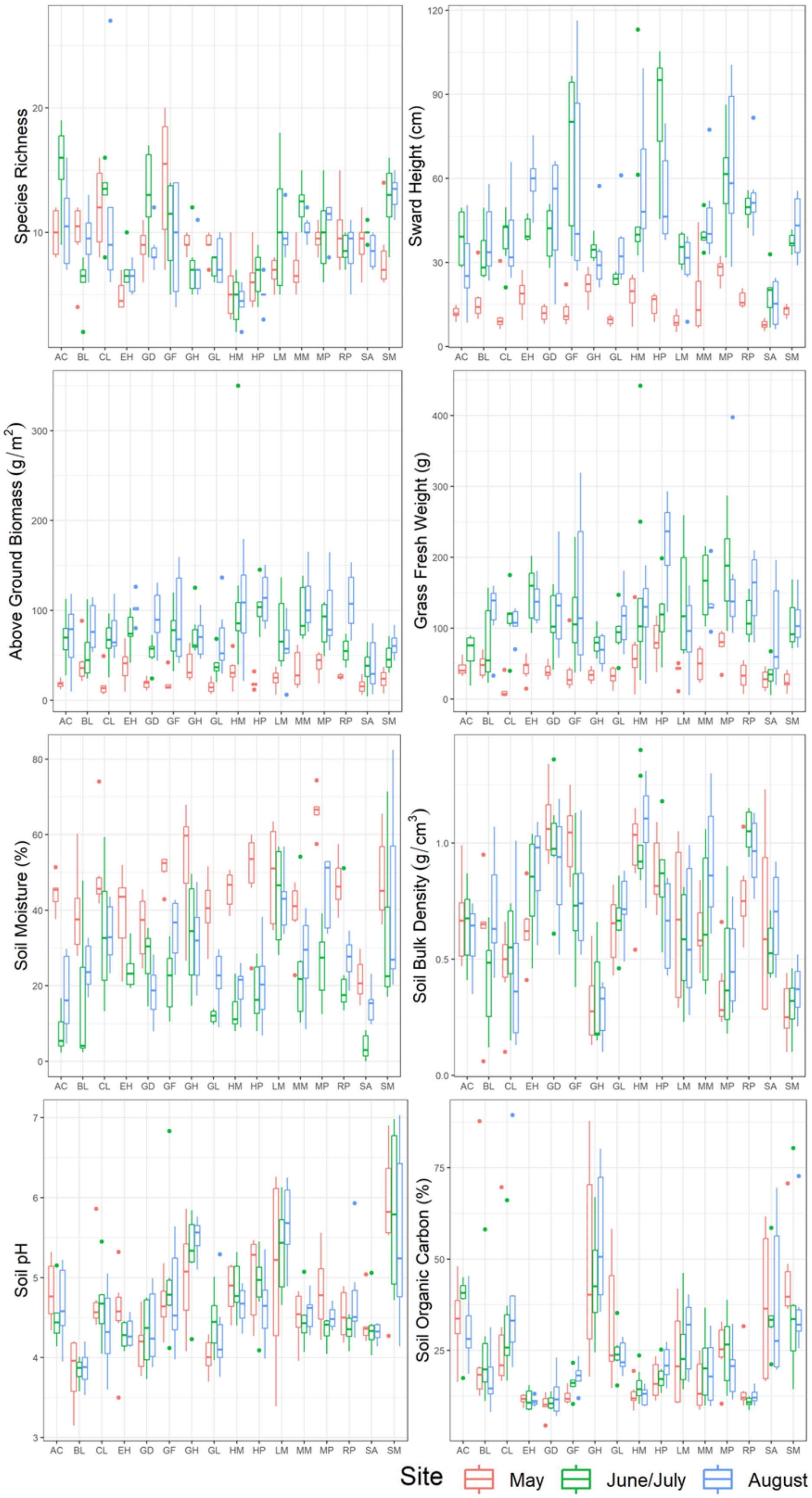


Figure 4-8. Variation in grassland environmental traits across 16 species-rich grasslands in Scotland. Species richness (number of species), above ground biomass (g/m^2), grass fresh weight (g), average sward height (cm), average soil moisture (%), average soil bulk density (g/cm^3), average soil organic carbon content (%), and average soil pH.

4.4.3 Classification Schema for Species-Rich Grasslands

The final broad classification schema, used throughout the thesis for the habitat classification model, was adapted from the UK BAP phase 1 broad habitats to include a further SRG category of coastal grasslands. This is because the species presence found *in situ* was markedly different than the other SRG classes, with high dominance of certain coastal species, such as *Armeria maritima* and *Silene uniflora*. Classification schemes exclude coastal grasslands as an SRG class, however, they are areas of high species richness that are utilised by grassland invertebrate species, such as *A. Artaxerxes* and have their own importance for identification. The broad habitats are previously based off their typical soil pH but when investigating the feature importance in the classification it was found that a) feature importance is altered by the classification system used, and b) pH is a low predictor of SRG type compared to other environmental and topographical variables. The five classes spectral profiles also demonstrate that a coastal SRG has a markedly different spectral profile, with lower reflectance values seen across the red-edge bands (Figure 4-9).

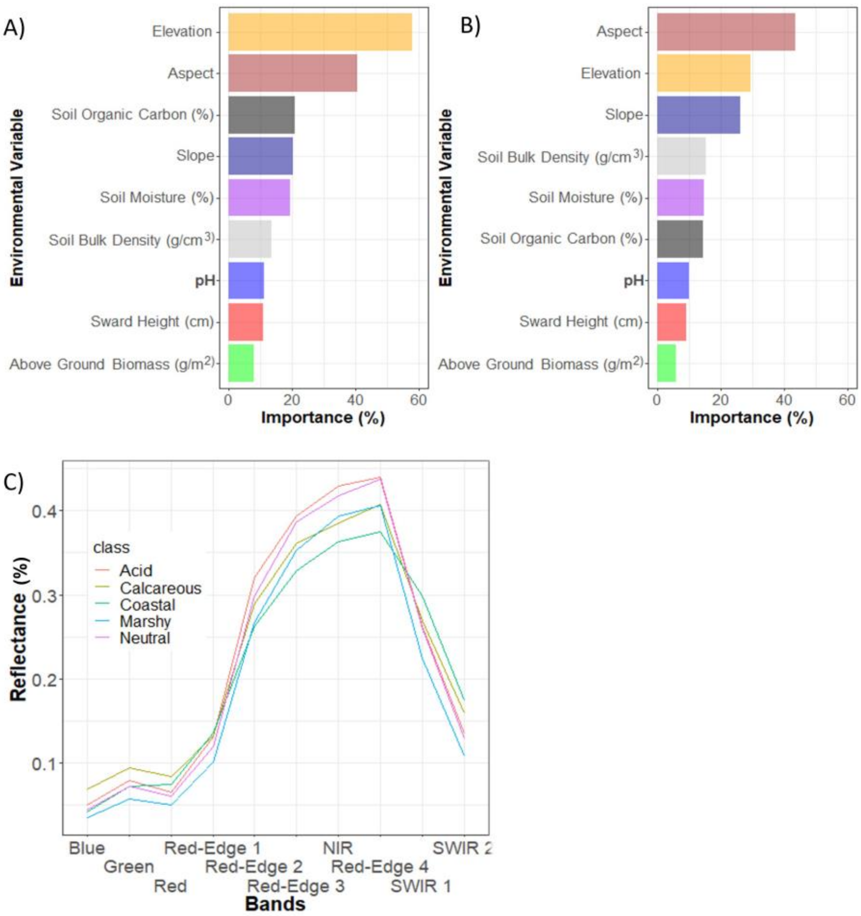


Figure 4-9. Environmental feature importance for grassland class prediction. A) Europe's EUNIS habitat classification, B) UK broad habitat classification. Importance of the feature in habitat class prediction is measured by the mean decrease in the Gini coefficient (%), where the greater the decrease, the greater the importance of a variable. C) Spectral signatures (reflectance %) of five grassland classes across Sentinel-2 bands.

4.4.4 Model Outputs and Accuracy Assessment

Training of RF and SVM models consisted of 10 cross-validations, repeated 3 times. Model accuracies were very similar between SVM and RF (Table 4-10). Although SVM slightly outperformed RF, RF was the final classifier chosen as it is more suitable for multi-class classifications (Alshari and Gawali, 2022). Not only this but RF tends to be more stable and is less time consuming, which is more applicable for national scale mapping (Belgiu and Drăguţ, 2016; Rodriguez-Galiano *et al.*, 2012). The SVM overall accuracy of 1 also suggests that the model is overfitting and there were scaling issues found with using the SVM method.

Table 4-10. Random Forest and Support Vector Machine average classification accuracies of training data.

Model	Overall Accuracy	Kappa
RF	0.986	0.983
SVM	1	1

The final RF model was tuned to the highest accuracy (mtry = 2), resulting in an Out-of-Bag error of 1.42%. Classification errors were seen in the acid, calcareous, and neutral SRG types in the training data. However, when the model was applied to the unseen test data, coastal grasslands were the worst predicted, with misclassifications seen also in neutral SRGs (Table 4-11).

Table 4-11. Classification errors in training data (top) and balanced accuracies in testing data (bottom) per class in final Random Forest habitat classification model used for species-rich grassland prediction across Scotland.

Class	Acid	Artificial Surface	Bare Soil	Calcareous	Improved	Marshy	Neutral	Coastal	Water	Wood-land	Error (%)
Acid	14						3				0.176
Artificial Surface		95			1						0.010
Bare Soil			83								0.000
Calcareous		1		4							0.200
Improved					88						0.000
Marshy						5					0.000
Neutral	1						37				0.026
Coastal								5			0.000
Water									56		0.000
Woodland					1					98	0.010

Class	Acid	Artificial Surface	Bare Soil	Calcareous	Improved	Marshy	Neutral	Coastal	Water	Wood-land	Balanced Accuracy (%)
Acid	7										1.000
Artificial Surface		40	1								0.997
Bare Soil			34								0.986
Calcareous				1							1.000
Improved					37						1.000
Marshy						1					1.000
Neutral							15	1			0.997
Coastal											0.500
Water									24		1.000
Woodland										42	1.000

The model showed that the features of most importance in the classifications were the VIs (specifically NDVI and GVI), whereas band importance varied a lot by season. Aspect and soil texture show consistently lower importance in predictions (Appendix C-3).

4.5 Discussion

The research in this chapter resulted in the selection of 16 SRG locations across Scotland. Selected sites, varied in their SRG class and abiotic conditions, and were chosen to incorporate wider SRG variations, as well as to include specific habitat for *A. artaxerxes*. The 16 sites were visited in summer 2021 and environmental data collection occurred to confirm the SRG class of each specific site. Species richness varied from 2 to 27 species across the quadrats. The community sward heights ranged from 5.2 cm to 116.2 cm, whilst AGB ranged from 4.4 g/m^2 to 349.8 g/m^2. Soil moisture content in a quadrat varied from 0% to 86.6%. The pH of the soil in a quadrat ranged over 3.2 to 7, whilst the organic content varied 4.4% to 89.4%. Finally, the soil bulk density across the quadrats ranged from 0.06 g/cm^3 to 1.4 g/cm^3. RS data was also collected for each site.

This data, along with the resulting habitat classifications were used to test multiple classification models, resulting in a final classification model that predicted SRGs with a high accuracy of 98.6%. This model was then applied to subsequently acquired S2 satellite imagery to nationally predict SRGs. These prediction maps were used to create the outputs that are associated with the created citizen science survey, all of which is further discussed in chapter 6. The results and considerations here are in relation to the process of locating potential SRG sites and the finalised habitat classification model.

4.5.1 Data Challenges for Locating Species-Rich Grassland Sites

A range of climatic and environmental conditions, in relation to *A. artaxerxes* distribution, were mapped and aligned with the potential areas of SRGs across Scotland. From this, 25 potential site locations were identified. Due to time commitments and permissions, 16 were shortlisted as final potential SRG sites. Figure 4-6 illustrates a significant difference in coverage of potential areas of SRG from the various sources. The habitat features of HabMos are more specific to SRGs, as were mapped using NVC surveys, allowing different communities that correspond to certain EUNIS codes to be extracted at a hierarchical level. However, as the dataset is incomplete there are large gaps in coverage. For example, on the east coast of Scotland the occurrence of *A. artaxerxes* is high, however, there is little overlap between the potential SRG areas as identified by the data from NatureScot HabMoS, even though it is known that this butterfly is found largely on calcareous SRGs. There were certain NVC communities that were included as a broader habitat EUNIS classification in the extracted dataset, potentially as they had not been identified to a higher level. For example, there are features of the broad habitat of “Moist or wet eutrophic and mesotrophic grassland” (E3.4), which could contain either species-rich wet grasslands “[*Juncus acutiflorus*] meadows” (E4.42) or the “Flood swards and related communities” (E3.44), which are not a species-rich grassland classification.

Therefore, there is some inclusion of broader habitats that may not strictly consist of just species-rich grasslands, as it was not possible to extract information at a higher resolution.

Similarly, the UK CEH land cover data set was limited to only the broad classifications. Therefore, more detailed communities could not be extracted to remove communities of lower species-richness. For example, those in acid grasslands habitat E1.7, including species-poor *Nardus* grasslands, or marsh grasslands that are truly fens or swamps. This denotes that this feature map includes areas of both potential SRG and otherwise semi-natural grasslands that are also species-poor. There were some limitations identified by the UK CEH in this dataset, for example, certain habitat features were unable to be identified through the satellite imagery chosen, and manual corrections were not completed due to time restrictions. This resulted in accuracy levels of 79.4% for habitat classifications for this data (Morton *et al.*, 2020). The lowland grassland database consisted of surveys dating back between the 1980-90s and do not include the full extent of SRGs in Scotland either (Pers.comms., 2020). It is also possible that these areas may now have been converted to new land uses over the past few decades. However, the layer was still useful to try and locate SRGs of conservation interest, for both priority habitats and protected species, as they have been determined as SRGs.

It must be noted that the HabMoS and UK CEH land cover maps were not created for the purposes of identifying areas of SRG. The use of the UK CEH map is to address land cover changes and, as such, it may not be appropriate to use this map solely for the identification of some SRG sites in Scotland. Whilst the HabMoS map is unfinished and identifies habitats based on EUNIS classifications (of which SRG is not a discrete classification). In this instance, both datasets were needed with the inclusion of other data, such as the lowland grassland database and indicator species occurrence records, to locate areas of potential SRGs for *in situ* measurements.

No independent layer would have been solely suitable for the identification of SRGs due to the various caveats with each. The combination of the layers was most beneficial to try account for any limitations within a layer that would impact the site selection. However, the exploration of these layers further highlights the need for this work to be conducted. It is necessary to create a tool that will be able to gain the coverage, as well as detail, of SRGs in Scotland, to map these priority habitats for their conservation and their associated species.

4.5.2 Considerations of the Classification Schema

Once the sites had been initially visited in 2021 and determined suitable for SRG characterisation, the classification schema was refined from the literature. The process of determining each site and quadrat classification was not simple. Grassland broad habitats are often determined by their soil pH, as outlined previously. However, other methods of classification look at species presence, which are then used as indicator species for a particular broad grassland classification. Findings from the field campaigns showed that although there was a range of soil pH on sites, most averages were indicating acidic soils. This, however, did not necessarily reflect the plant communities that were found on the sites. Many sites had a high number of species that would usually be found on more neutral soils (e.g., *Holcus lanatus, Dactylis glomerata, Cirsium arvense*). Even species that are usually found on calcareous soils (e.g., *H. nummularium, Galium verum, Succisa pratensis*) were being found on soils with low pH (Nature Conservancy Council, 1990). This would suggest a different classification depending on whether the habitat was classified based on soil pH or plant community.

There are several reasons as to why the soil pH may have been lower than expected. For one, the method of determining the soil pH involved the use of calcium chloride (a standard used in measuring soil pH). Calcium chloride does lower the pH but not usually enough to change the pH index. Not only this, but soil samples were extracted from the top layer (also a field standard) which is found to have a higher organic content, resulting in a lower pH. However, the pHs of the first field campaign were tested via three methods (strip test, in field pH probe, and lab probe analysis post soil drying) (Table 4-12) and, although the average soil pH of field tests was slightly higher in some (but not all) cases, this is seen in other pH comparisons and could be due to the calcium chloride content, or the less standardized conditions of in-field tests. Storing samples for pH analysis also does not appear to be problematic (Chou et al., 2016; Pansu and Gautheyrou, 2002).

Table 4-12. Site comparison of soil pHs measured with a field probe versus with a lab probe.

Site	Average Field Soil pH	Average Lab Soil pH
AC	5.61	4.82
BL	4.11	3.66
CL	5.00	4.77
EH	4.87	4.54
GD	4.64	4.63
GF	4.86	4.88
GH	4.93	4.29
GL	4.61	4.66
HM	5.82	4.88
HR	4.90	5.15
MM	5.04	4.46
MP	4.75	4.84
RP	5.50	4.54

Instead, what can explain the presence of a wide range of species on acidic soils can be high tolerances of most species. Plant communities or individual species have been found to grow and even thrive on a range of soil pH; for example, *H. nummularium*, a calcareous loving wildflower can still grow on mildly acidic soils (Gallacher, n.d.; Mill, n.d.). This suggests that when a combination of conditions can satisfy the needs of species, they may be found where unexpected. The underlying bed rock and soil depth can also have a considerable influence, allowing various species to thrive in unassuming places. For example, calcareous upland grasslands have indicated acidic conditions where the lime has been leached and, as such, acidic species like common tormentil become present (The

Wildlife Trusts, 2021). This means that the habitat classification process may become more difficult, as communities do not fit neatly into boxes.

When analysing the feature importance of SRG classification with different classification schemas (e.g., EUNIS versus UK BAP phase 1), this altered which features were most important in the classification designation. This highlights that the chosen schema will alter the final classification of a site, even if the species presence is the same. The results also suggest that pH is a low determinant in multiple classification schemas of SRGs and should not be used as the sole predictor of these classes where other variables had a greater importance. The classification procedure was then based on species presence to indicate specific broad/Phase 1 habitats with associated descriptions.

The different methods of classification were not designed interchangeably and, as such, comparisons between schemes are not straightforward. This can explain why rush pastures and purple moor grass habitats can be associated with bogs, wet heaths, mires, and marshy grasslands depending on the classification system (Botanaeco, 2020). Even within classification schema, such as the Phase 1 and NVC surveys, there is a lot of variation, with many NVC communities being found across different Phase 1 classifications (e.g., a calcareous NVC community such as CG10 can be associated with both acid and calcareous soils). The UK Habitat Classification Working Group are working towards a single classification system; however, its wide scale use is unknown (UKHab, 2023).

Due to these considerations, the final classification schema that was used in the remote sensing model creation was altered from schemas based in literature. The nomenclature of the class names was kept the same (e.g., calcareous, acid, neutral) for general understanding and alignment but were adapted to be based purely off species presence of neutral/calcareous/acid/marshy loving species rather than the pH of the soil. Because of this, a 5th broad SRG type was included – coastal – based off the very specific coastal species presence that were not found in other SRG habitats. When out surveying, visual classification is the most feasible for the public to conduct due to lack of equipment needed to measure variables, such as soil pH, making this a more accessible method, especially if accompanied by trained professionals or with the help of guidebooks and handouts. Basing the classifications on species presence allows class determination to be simpler for participants of the citizen science survey ground-truthing that followed the habitat prediction modelling.

4.5.3 Model Accuracy and Limitations

The model shows a high classification accuracy across the classes, even between the multiple grassland types. It is difficult to compare the model results to published studies, as no classification research has been conducted specifically on this thesis' defined grassland classes in the UK. Other studies that have attempted semi-natural grassland classification mapping have used varying classification schemas, highlighting the need for a more synergistic system. However, as some of these studies were also conducted in other countries, it further hinders comparison, as communities (and their associated descriptions) vary across nations. This raises the question of how wide-scale, global mapping of grassland communities, at this breadth, could be feasible. However, it is important to investigate if semi-natural grassland mapping is possible and then further research the application of the tool in different contexts. Where semi-natural grassland mapping has occurred, initial model overall accuracies have been found between 71% – 91.1% (Raab *et al.,* 2018; Rapinel *et al.,* 2019; Schuster *et al.,* 2015; Zlinsky *et al.,* 2014). These accuracies are high but may be slightly lower than the accuracies reported here due to the scale of the communities measured and satellite imagery used. The mapping here considers multiple communities contained within the broader SRG classes, rather than to NVC level which is more comparable to these studies.

It must be noted that the accuracies described in the results are based on train and test data, and ground-truthing to ascertain real-world accuracies were conducted for subsequent chapters (chapter 6), utilising the designed citizen science survey as part of this thesis. There are certain limitations that must be considered when interpreting the model results. The model included a total of 10 habitat classes (five of which were SRG). It was not possible to include all defined natural habitats in the model training data, such as the inclusion of wetlands and moorlands, due to time restrictions of data collection and available habitat maps online. As the model must classify every pixel in an image it is likely that there are overestimates of certain classes. For example, some moorland habitats are likely to be included in the predicted acidic grassland areas, whilst some wetlands may be included in the predicted marshy grassland areas. This is due to similar spectral profiles that might be seen between the habitats. Some areas of SRGs may be underpredicted due to transitional habitats. Where scrub tends to be present on SRG sites, this could more greatly influence the spectral profile of a pixel due to the larger spatial scale of scrub and, therefore, classify certain pixels as woodland.

Unfortunately, due to the distribution of SRG classes across Scotland, the number of data points for the SRG types included in the model could not be equal. For example, only one calcareous, one marshy, and one coastal SRG site were able to be visited. In comparison, there were 9 neutral SRG sites. However, classification models should not be fixed, and more data can be added as further research is conducted. This will allow models to continuously be updated with the most relevant information and continue to improve accuracies on habitat mapping.

4.6 Conclusions and Next Steps

National mapping of species-rich grasslands has been identified as an organisational and governmental goal in policy. This has not succeeded past local attempts. Data collation of potential and known SRG locations across Scotland is found in published open-source data, but updates and widening this reach is needed. The analysis of open-source secondary data outlined here allowed SRG sites to be located across Scotland for their characterisation and classification. It is evident that numerous classification schemes and naming conventions cause confusion in both national and international mapping attempts, and it is suggested future research should work across governing bodies to unite these techniques, as the EUNIS and UKHab schemes have started to attempt. The model created here initially endeavours to discern multiple SRG classes from remote sensing, an effort that has had little attention in the past. The high accuracies reported suggest that there is potential for the model to successfully send participants of the citizen science survey to new areas of SRGs in Scotland, who will confirm how well the model has performed.

In situ surveying continued for a second-year field campaign in 2022, following the methodology and refinement of identified sites highlighted in this chapter. This data was used to try enhancing SRG mapping through common techniques employed in RS science. Multiple RS devices across spatial and spectral resolutions were tested in their ability to retrieve and predict SRG traits and improve the habitat mapping through the addition of the acquired information, if successful. These results are found in the subsequent chapter 5. The methods associated with the extraction of ground-truthing locations are detailed in chapter 6 of this thesis, along with the associated results. The research here is one of the first to apply remote sensing to species-rich grassland classification at this scale.

References

Adankon M., Cheriet M., 2009. Support Vector Machine. In: Li S.Z., Jain A. (eds) *Encyclopedia of Biometrics*. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-73003-5_299

Al-Doski, J., Mansorl, S.B. and Shafri, H.Z.M., 2013. Image classification in remote sensing. *Department of Civil Engineering, Faculty of Engineering, University Putra, Malaysia*.

Ali, I., Cawkwell, F., Dwyer, E., Barrett, B. and Green, S., 2016. Satellite remote sensing of grasslands: from observation to management. *Journal of Plant Ecology*, *9*(6), pp.649-671.

Alshari, E.A. and Gawali, B.W., 2022. Analysis of Machine Learning Techniques for Sentinel-2A Satellite Images. *Journal of Electrical and Computer Engineering*, *2022*.

Badreldin, N., Prieto, B. and Fisher, R., 2021. Mapping grasslands in Mixed Grassland ecoregion of Saskatchewan using big remote sensing data and machine learning. *Remote Sensing*, *13*(24), p.4972.

Bekkema, M.E. and Eleveld, M., 2018. Mapping grassland management intensity using sentinel-2 satellite data. *GI_Forum*, *2018*(1), pp.194-213.

Belgiu, M. and Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, pp.24-31.

Boegh, E., Houborg, R., Bienkowski, J., Braban, C.F., Dalgaard, T., Dijk, N.V., Dragosits, U., Holmes, E., Magliulo, V., Schelde, K. and Tommasi, P.D., 2013. Remote sensing of LAI, chlorophyll and leaf nitrogen pools of crop-and grasslands in five European landscapes. *Biogeosciences*, *10*(10), pp.6279-6307.

Boori, M.S., Paringer, R., Choudhary, K. and Kupriyanov, A., 2018, August. Supervised and unsupervised classification for obtaining land use/cover classes from hyperspectral and multi-spectral imagery. In *Sixth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2018)* (Vol. 10773, p. 107730L). International Society for Optics and Photonics.

Botanaeco., 2020. Fitting Phase 1 habitats & NVC communities to their designations. [Online]. Available at: https://botanaeco.co.uk/blog/f/fitting-phase-1-habitats-nvc-communities-to-their-designation. Accessed 19th May 2023.

Bourne, A., 2020. Species-rich grasslands. *[Online]*. Accessed 21st December 2020. Available at: https://storymaps.arcgis.com/stories/6b086586a4194d32992ee2564d6979a0

Boyd, D.S., Foody, G.M., Brown, C., Mazumdar, S., Marshall, H. and Wardlaw, J., 2022. Citizen science for Earth Observation (Citizens4EO): understanding current use in the UK. *International Journal of Remote Sensing*, *43*(8), pp.2965-2985.

Brereton, T., Moss, D., Greatorex-Davies, N., Hobson, R., and Feber, R., 2000. Habitat classification for butterfly transects. *[Online]*. Accessed 10th December 2020. Available at: https://www.ukbms.org/Downloads/UKBMS%20G4%20Transect%20habitat%20classification.pdf

Buck, O., Millán, V.E.G., Klink, A. and Pakzad, K., 2015. Using information layers for mapping grassland habitat distribution at local to regional scales. *International Journal of Applied Earth Observation and Geoinformation*, *37*, pp.83-89.

Byrne, D., Moran, J., Maher, C. and Jones, G., 2018. Measure Handbook: Results-based Agri-environment Pilot Schemes – Species-rich Grassland and Wet Grasslands and Mosaic Habitat Suitable for Marsh Fritillary. Report prepared for the European Union, Agreement No. 07.027722/2014/697042/SUB/B2

Ceballos, G., Ehrlich, P.R. and Dirzo, R., 2017. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the NATIONAL Academy of Sciences*, *114*(30), pp.E6089-E6096.

Chitale, V.S., Behera, M.D. and Roy, P.S., 2019. Deciphering plant richness using satellite remote sensing: a study from three biodiversity hotspots. *Biodiversity and Conservation*, *28*(8), pp.2183-2196.

Chou, W.C., Gong, G.C., Yang, C.Y. and Chuang, K.Y., 2016. A comparison between field and laboratory pH measurements for seawater on the E ast C hina S ea shelf. *Limnology and Oceanography: Methods*, *14*(5), pp.315-322.

Chytrý, M., Tichý, L., Hennekens, S.M., Knollová, I., Janssen, J.A., Rodwell, J.S., Peterka, T., Marcenò, C., Landucci, F., Danihelka, J. and Hájek, M., 2020. EUNIS Habitat Classification: Expert system, characteristic species combinations and distribution maps of European habitats. *Applied Vegetation Science*, *23*(4), pp.648-675.

Clevers, J.G.P.W. and Gitelson, A.A., 2012, April. Using the red-edge bands on Sentinel-2 for retrieving canopy chlorophyll and nitrogen content. In *First Sentinel2: Preparatory symposium* (pp. 23-27).

Crofts, A. and Jefferson, R.G., 1999. Grassland management in wood-pasture and parkland. *The lowland grassland management handbook. 2nd edition. English Nature, The Wildlife Trusts*.

Dadds, N.J. and Averis, A.B.G. 2014. The extent and condition of non-designated species rich lowland grasslands in Scotland. Scottish Natural Heritage Commissioned Report No. 571.

Dahlström, A., Iuga, A.M. and Lennartsson, T., 2013. Managing biodiversity rich hay meadows in the EU: a comparison of Swedish and Romanian grasslands. *Environmental Conservation*, *40*(2), pp.194-205.

De Simone, W., Allegrezza, M., Frattaroli, A.R., Montecchiari, S., Tesei, G., Zuccarello, V. and Di Musciano, M., 2021. From Remote Sensing to Species Distribution Modelling: An Integrated Workflow to Monitor Spreading Species in Key Grassland Habitats. *Remote Sensing*, *13*(10), p.1904.

Divíšek, J. and Chytrý, M., 2018. High-resolution and large-extent mapping of plant species richness using vegetation-plot databases. *Ecological Indicators*, *89*, pp.840-851.

Dlamini, S., Nhapi, I., Gumindoga, W., Nhiwatiwa, T. and Dube, T., 2016. Assessing the feasibility of integrating remote sensing and in-situ measurements in monitoring water quality status of Lake Chivero, Zimbabwe. *Physics and Chemistry of the Earth, Parts A/B/C*, *93*, pp.2-11.

Dudley, N., Eufemia, L., Fleckenstein, M., Periago, M.E., Petersen, I. and Timmers, J.F., 2020. Grasslands and savannahs in the UN Decade on Ecosystem Restoration. *Restoration Ecology*, *28*(6), pp.1313-1317.

El-Rawy, M., Fathi, H. and Abdalla, F., 2019. Integration of remote sensing data and in situ measurements to monitor the water quality of the Ismailia Canal., Nile Delta, Egypt. *Environmental Geochemistry and Health*, pp.1-20.

Enderle, D.I. and Weih Jr, R.C., 2005. Integrating supervised and unsupervised classification methods to develop a more accurate land cover classification. *Journal of the Arkansas Academy of Science*, *59*(1), pp.65-73.

Erdanaev, E., Kappas, M. and Wyss, D., 2022. The Identification of Irrigated Crop Types Using Support Vector Machine, Random Forest and Maximum Likelihood Classification Methods with Sentinel-2 Data in 2018: Tashkent Province, Uzbekistan. *International Journal of Geoinformatics*, *18*(2).

ESA., 2013. User Handbook. *ESA Standard Document*, *64*.

Esri Inc., 2020. *ArcGIS Pro* (Version 2.5). Esri Inc. https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview.

Evans, D., 2006, November. The habitats of the European Union habitats directive. In *Biology and Environment: Proceedings of the Royal Irish Academy* (pp. 167-173). Royal Irish Academy.

Evans, D., 2010. Interpreting the habitats of Annex I: past, present and future. *Acta Botanica Gallica*, *157*(4), pp.677-686.

Evans, D., 2012. Building the European union's Natura 2000 network. *Nature Conservation*, *1*, pp.11-26.

Fassnacht, F.E., Müllerová, J., Conti, L., Malavasi, M. and Schmidtlein, S., 2022. About the link between biodiversity and spectral variation. *Applied Vegetation Science*, *25*(1), p.e12643.

Fox R, Dennis EB, Purdy KM, Middlebrook I, Roy DB, Noble DG, Botham MS & Bourn NAD.,2023. The State of the UK's Butterflies 2022. Butterfly Conservation, Wareham, UK

Franke, J., Keuck, V. and Siegert, F., 2012. Assessment of grassland use intensity by remote sensing to support conservation schemes. *Journal for Nature Conservation*, *20*(3), pp.125-134.

Gallacher, J., 2007. The northern brown argus in the Ochil hills, Clackmannanshire. *Forth Naturalist and Historian*, *30*, p.73.

Gholizadeh, H., Gamon, J.A., Helzer, C.J. and Cavender-Bares, J., 2020. Multi-temporal assessment of grassland α-and β-diversity using hyperspectral imaging. *Ecological Applications*, *30*(7), p.e02145.

Gholizadeh, H., Gamon, J.A., Townsend, P.A., Zygielbaum, A.I., Helzer, C.J., Hmimina, G.Y., Yu, R., Moore, R.M., Schweiger, A.K. and Cavender-Bares, J., 2019. Detecting prairie biodiversity with airborne remote sensing. *Remote Sensing of Environment*, *221*, pp.38-49.

Gov Scot., 2018. Habitat map of Scotland. *[Online]*. Accessed 21st December 2020. Available at: https://www.environment.gov.scot/our-environment/habitats-and-species/habitat-map-of-scotland/#what

Hijmans, R., 2023. *raster: Geographic Data Analysis and Modeling*. R package version 3.6-21, https://rspatial.org/raster.

Ichter, J., Evans, D. and Richard, D., 2014. Terrestrial habitat mapping in Europe: an overview. *Publications Office of the European Union, Luxembourg*.

Imran, H.A., Gianelle, D., Rocchini, D., Dalponte, M., Martín, M.P., Sakowska, K., Wohlfahrt, G. and Vescovo, L., 2020. VIS-NIR, red-edge and NIR-shoulder based normalized vegetation indices response to co-varying leaf and Canopy structural traits in heterogeneous grasslands. *Remote Sensing*, *12*(14), p.2254.

Imran, H.A., Gianelle, D., Scotton, M., Rocchini, D., Dalponte, M., Macolino, S., Sakowska, K., Pornaro, C. and Vescovo, L., 2021. Potential and limitations of grasslands α-diversity prediction using fine-scale hyperspectral imagery. *Remote Sensing*, *13*(14), p.2649.

Jackson, D.L. 2000. Guidance on the interpretation of the Biodiversity Broad Habitat Classification (terrestrial and freshwater types): Definitions and the relationship with other classifications, JNCC Report No. 307, JNCC, Peterborough, ISSN 0963-8091.

Jefferson, R.G., Smith, S.L.N. & MacKintosh, E.J. 2019. Guidelines for the Selection of Biological SSSIs. Part 2: Detailed Guidelines for Habitats and Species Groups. Chapter 3 Lowland Grasslands. Joint Nature Conservation Committee, Peterborough.

JNCC., 2010. *Handbook for Phase 1 Habitat Survey – a Technique for Environmental Audit*. Joint Nature Conservation Committee, Peterborough

JNCC., 2019a. Terrestrial habitat classification schemes. *[Online]*. Accessed on 10th December 2020. Available at: https://jncc.gov.uk/our-work/terrestrial-habitat-classification-schemes/

JNCC., 2019b. UK BAP priority habitats. *[Online]*. Accessed on 7th December 2020. Available at: https://jncc.gov.uk/our-work/uk-bap-priority-habitats/#uk-bap-broad-habitats

Karagiannopoulou, A., Tsertou, A., Tsimiklis, G. and Amditis, A., 2022. Data fusion in earth observation and the role of citizen as a sensor: a scoping review of applications, methods and future trends. *Remote Sensing*, *14*(5), p.1263.

King, M., 2010, December. An investigation into policies affecting Europe's semi-natural grasslands. In *A report by The Grasslands Trust commissioned by the European Forum on Nature Conservation and Pastoralism,(1097893)* (Vol. 35).

Kral., K.C.A., 2018. Method Selection and Estimating Densities for Grassland Butterflies in the Mixed-Grass Prairie. Accessed 12th March 2021. Available at: https://library.ndsu.edu/ir/bitstream/handle/10365/27855/Method%20Selection%20and%20Estimating%20Densities%20for%20Grassland%20Butterflies%20in%20the%20Mixed-Grass%20Prairie.pdf?sequence=3&isAllowed=y

Li, C., Zhou, L. and Xu, W., 2021. Estimating aboveground biomass using Sentinel-2 MSI data and ensemble algorithms for grassland in the Shengjin Lake Wetland, China. *Remote Sensing*, *13*(8), p.1595.

Manning, P., de Vries, F.T., Tallowin, J.R., Smith, R., Mortimer, S.R., Pilgrim, E.S., Harrison, K.A., Wright, D.G., Quirk, H., Benson, J. and Shipley, B., 2015. Simple measures of climate, soil properties and plant traits predict national‐scale grassland soil carbon stocks. *Journal of Applied Ecology*, *52*(5), pp.1188-1196.

Meng, B., Zhang, Y., Yang, Z., Lv, Y., Chen, J., Li, M., Sun, Y., Zhang, H., Yu, H., Zhang, J. and Lian, J., 2022. Mapping grassland classes using unmanned aerial vehicle and MODIS NDVI data for temperate grassland in inner Mongolia, China. *Remote Sensing*, *14*(9), p.2094.

Michalcová, D., Chytrý, M., Pechanec, V., Hájek, O., Jongepier, J.W., Danihelka, J., Grulich, V., Šumberová, K., Preislerová, Z., Ghisla, A. and Bacaro, G., 2014. High plant diversity of grasslands in a landscape context: a comparison of contrasting regions in Central Europe. *Folia Geobotanica*, *49*(2), pp.117-135.

Mill. L., n.d. *Helianthemum nummularium. [Online]*. Available at: https://pfaf.org/user/Plant.aspx?LatinName=Helianthemum+nummularium. Accessed 18th May 2023.

Moon, M., Richardson, A.D. and Friedl, M.A., 2021. Multiscale assessment of land surface phenology from harmonized Landsat 8 and Sentinel-2, PlanetScope, and PhenoCam imagery. *Remote Sensing of Environment*, *266*, p.112716.

Morton, R. D.; Marston, C.G.; O'Neil, A. W.; Rowland, C. S., 2020. Land Cover Map 2019 (25m rasterised land parcels, GB). NERC Environmental Information Data Centre. (Dataset). https://doi.org/10.5285/f15289da-6424-4a5e-bd92-48c4d9c830cc

Moss, D., 2014. EUNIS habitat classification–a guide for users. *European Topic Centre on Biological Diversity*.

Muro, J., Linstädter, A., Magdon, P., Wöllauer, S., Männer, F.A., Schwarz, L.M., Ghazaryan, G., Schultz, J., Malenovský, Z. and Dubovyk, O., 2022. Predicting plant biomass and species richness in temperate grasslands across regions, time, and land management with remote sensing and deep learning. *Remote Sensing of Environment*, *282*, p.113262.

Natural England., 2008. Soil texture (TIN037). *[Online]*. Accessed 23rd February 2021. Available at: turalengland.org.uk/publication/32016#:~:text=Texture%20describes%20the%20mixture%20of,indicate%20their%20ease%20of%20cultivation.

Nature Conservancy Council., Britain, G. and Unit, E.F., 1990. *Handbook for Phase 1 habitat survey: a technique for environmental audit*. Nature Conservancy Council.

Nature Scot., 2021. Species rich grasslands guidance leaflet. [Online]. Accessed 19th June 2023. Available at: https://www.nature.scot/sites/default/files/2021-01/Species%20Rich%20Grasslands%20guidance%20leaflet.pdf

Nevis Landscape Partnership., 2015. Advice on habitat suitability and monitoring approaches for Lepidoptera in mid and upper Glen Nevis. [Online]. Accessed 12th March 2021. Available at: Microsoft Word - NLP Lepidoptera habitat survey 2015 REPORT_FINAL.doc (nevislandscape.co.uk)

Palmer, M.W., Wohlgemuth, T., Earls, P., Are Âvalo, J.R., Thompson, S.D., 2000. Opportunities for long-term ecological research at the Tallgrass Prairie Preserve, Oklahoma. Pages 123±128. In *Cooperation in Long Term Ecological Research in Central and Eastern Europe: Proceedings of the ILTER Regional Workshop,* Lajtha K, Vanderbilt K (eds). 22±25 June, 1999, Budapest, Hungary. Oregon State University: Corvallis, OR.

Pansu, M. and Gautheyrou, J., 2006. pH measurement. *Handbook of Soil Analysis: Mineralogical, Organic and Inorganic Methods*, pp.551-579.

Pause, M., Schweitzer, C., Rosenthal., M., Keuck, V., Bumberger, J., Dietrich, P., Heurich, M., Jung, A. and Lausch, A., 2016. In situ/remote sensing integration to assess forest health—A review. *Remote Sensing*, *8*(6), p.471.

Peciña, M.V., Bergamo, T.F., Ward, R.D., Joyce, C.B. and Sepp, K., 2021. A novel UAV-based approach for biomass prediction and grassland structure assessment in coastal meadows. *Ecological Indicators*, *122*, p.107227.

Pellet, J., Bried, J.T., Parietti, D., Gander, A., Heer, P.O., Cherix, D. and Arlettaz, R., 2012. Monitoring butterfly abundance: beyond Pollard walks. *PloS One*, *7*(7), p.e41396.

Peng, Y., Fan, M., Bai, L., Sang, W., Feng, J., Zhao, Z. and Tao, Z., 2019. Identification of the best hyperspectral indices in estimating plant species richness in sandy grasslands. *Remote Sensing*, *11*(5), p.588.

Pigott, C.D., Ratcliffe, D.A., Malloch, A.J.C., Birks, H.J.B. and Proctor, M.C.F., 2000. *British Plant Communities: Volume 5, Maritime Communities and Vegetation of Open Habitats* (Vol. 5). Cambridge University Press.

Price, K.P., Guo, X. and Stiles, J.M., 2002. Optimal Landsat TM band combinations and vegetation indices for discrimination of six grassland types in eastern Kansas. *International Journal of Remote Sensing*, *23*(23), pp.5031-5042.

Psomas, A., Kneubühler, M., Huber, S., Itten, K. and Zimmermann, N.E., 2011. Hyperspectral remote sensing for estimating aboveground biomass and for exploring species richness patterns of grassland habitats. *International Journal of Remote Sensing*, *32*(24), pp.9007-9031.

Pywell, R.F., Bullock, J.M., Hopkins, A., Walker, K.J., Sparks, T.H., Burke, M.J. and Peel, S., 2002. Restoration of species-rich grassland on arable land: assessing the limiting processes using a multi-site experiment. *Journal of Applied Ecology*, *39*(2), pp.294-309.

Raab, C., Stroh, H.G., Tonn, B., Meißner, M., Rohwer, N., Balkenhol, N. and Isselstein, J., 2018. Mapping semi-natural grassland communities using multi-temporal RapidEye remote sensing data. *International Journal of Remote Sensing*, *39*(17), pp.5638-5659.

Rapinel, S., Mony, C., Lecoq, L., Clément, B., Thomas, A. and Hubert-Moy, L., 2019. Evaluation of Sentinel-2 time-series for mapping floodplain grassland plant communities. *Remote Sensing of Environment*, *223*, pp.115-129.

Reddy, C.S., 2021. Remote sensing of biodiversity: what to measure and monitor from space to species?. *Biodiversity and Conservation*, pp.1-15.

Reinermann, S., Asam, S. and Kuenzer, C., 2020. Remote sensing of grassland production and management—A review. *Remote Sensing*, *12*(12), p.1949.

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M. and Rigol-Sanchez, J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *67*, pp.93-104.

Rodwell, J. S., 1992. British Plant Communities: Grasslands and Montane Communities. Cambridge University Press, Cambridge.

Rossi, C., Kneubühler, M., Schütz, M., Schaepman, M.E., Haller, R.M. and Risch, A.C., 2022. Spatial resolution, spectral metrics and biomass are key aspects in estimating plant species richness from spectral diversity in species-rich grasslands. *Remote Sensing in Ecology and Conservation*, *8*(3), pp.297-314.

Schuster, C., Schmidt, T., Conrad, C., Kleinschmit, B. and Förster, M., 2015. Grassland habitat mapping by intra-annual time series analysis–Comparison of RapidEye and TerraSAR-X satellite data. *International Journal of Applied Earth Observation and Geoinformation*, *34*, pp.25-34.

Schwieder, M., Buddeberg, M., Kowalski, K., Pfoch, K., Bartsch, J., Bach, H., Pickert, J. and Hostert, P., 2020. Estimating Grassland Parameters from Sentinel-2: A Model Comparison Study. *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, *88*(5), pp.379-390.

Scobie, A. 2018. Habitat Map of Scotland Upland Mapping Pilot: developing a method to map upland habitats using stereo colour near-infrared aerial imagery. Scottish Natural Heritage Research Report No. 983.

Scottish Borders Council, date unknown. [Online]. Available at: https://www.scotborders.gov.uk/downloads/file/1372/grassland_and_enclosed_farmland. Accessed 7th March 2022.

Shellswell, C.H., Chant, J.J., Jefferson, R.G., Le Bas, B., Edwards, J. and Parton, C., 2016. Restoration of existing lowland grassland – timescales to achieve favourable condition. Plantlife, Salisbury. Accessed 25th January 2021. Available at: http://www.magnificentmeadows.org.uk/assets/pdfs/Lowland_grassland_timescales_for_recovery_advisory_note_FINAL-Design.pdf

Shen, M., Tang, Y., Klein, J., Zhang, P., Gu, S., Shimono, A. and Chen, J., 2008. Estimation of aboveground biomass using in situ hyperspectral measurements in five major grassland ecosystems on the Tibetan Plateau. *Journal of Plant Ecology*, *1*(4), pp.247-257.

Stenzel, S., Fassnacht, F.E., Mack, B. and Schmidtlein, S., 2017. Identification of high nature value grassland with remote sensing and minimal field data. *Ecological Indicators*, *74*, pp.28-38.

The Wildlife Trusts., 2021. Common Rock-rose. *[Online].* Accessed 17th February 2021. Available at: https://www.wildlifetrusts.org/wildlife-explorer/wildflowers/common-rock-rose

UK Biodiversity Action Plan., 2008. Priority Habitat Descriptions. BRIG (ed. Ant Maddock).

UKBMS., 2019. G1: monitoring butterfly numbers by the transect method - summary information for recorders. *[Online].* Accessed 4th March 2021. Available at: https://www.ukbms.org/Downloads/UKBMS%20G1%20Transect%20monitoring%20summary.pdf

UKHab., 2023. https://ukhab.org/. Accessed 18th May 2023.

Wachendorf, M., Fricke, T. and Möckel, T., 2018. Remote sensing as a tool to assess botanical composition, structure, quantity and quality of temperate grasslands. *Grass and Forage Science*, *73*(1), pp.1-14.

Yan, H., Xue, Z. and Niu, Z., 2021. Ecological restoration policy should pay more attention to the high productivity grasslands. *Ecological Indicators*, *129*, p.107938.

Yang, X., 2012. Assessing Responses of Grasslands to Grazing Management Using Remote Sensing Approaches. *University of Saskatchewan, Saskatoon.*

Yousefi, S., Mirzaee, S., Almohamad, H., Al Dughairi, A.A., Gomez, C., Siamian, N., Alrasheedi, M. and Abdo, H.G., 2022. Image classification and land cover mapping using sentinel-2 imagery: optimization of SVM parameters. *Land*, *11*(7), p.993.

Zhao, Y., Liu, Z. and Wu, J., 2020. Grassland ecosystem services: a systematic review of research advances and future directions. *Landscape Ecology*, pp.1-22.

Zhao, Y., Sun, Y., Chen, W., Zhao, Y., Liu, X. and Bai, Y., 2021. The potential of mapping grassland plant diversity with the links among spectral diversity, functional trait diversity, and species diversity. *Remote Sensing*, *13*(15), p.3034.

Zlinszky, A., Schroiff, A., Kania, A., Deák, B., Mücke, W., Vári, Á., Székely, B. and Pfeifer, N., 2014. Categorizing grassland vegetation with full-waveform airborne laser scanning: A feasibility study for detecting Natura 2000 habitat types. *Remote Sensing*, *6*(9), pp.8056-8087.

Chapter 5. Predicting Species Diversity and Community Traits from Remote Sensing in Species-Rich Grasslands

Authors Samantha Suter, Brian Barrett, and Natalie Welden conceived the ideas and designed methodology. Samantha Suter, Kenny Roberts, and Brian Barrett collected and processed the data. Samantha Suter analysed the data; Samantha Suter led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Status: Under review in *Remote Sensing Applications: Society and Environment* as of 25.04.24. Reformatted for thesis.

Abstract

Grassland mapping has been identified as a key conservation priority due to the essential ecosystem services provided by these habitats. Species-rich grasslands in Scotland are some of the least extensively mapped and most degraded habitats. Retrieval of grassland community traits from remote sensing is possible but has yet to be conducted widely across highly diverse species-rich grassland classes. We investigated whether species richness and grassland community traits (above ground biomass, sward height, fresh weight, and SPAD-measured chlorophyll-proxy) could be predicted across multiple spatial (8 cm – 10 m) and spectral (4 – 13 bands) scales using data acquired from Sentinel-2 and Planetscope satellites, and an Unoccupied Aerial Vehicle (UAV). From the results, we suggest a set of recommendations for remote sensing of species-rich grasslands. In contrast to studies of single sites and less diverse grasslands, we found that there was no significant relationship between the spectral diversity metrics; standard deviation and coefficient of variation, and species diversity (species richness) ($p = 0.211$ and $p = 0.141$ respectively) across seven of our study sites. Grassland trait prediction varied largely with spatial and spectral scale and the combination of predictor variables used. The UAV mounted Micasense predictor variables explained the most variance in predicting all traits: sward height, above ground biomass, fresh weight, and SPAD-measured chlorophyll-proxy ($R^2 = 0.545$, $R^2 = 0.221$, $R^2 = 0.235$, $R^2 = 0.167$ respectively). The results suggest that prediction estimates across multiple species-rich grassland classes using remote sensing may be hampered by increased variation and confounding factors in these highly diverse habitats. Further methodological advancements may be needed for wide scale cross-grassland habitat monitoring and mapping and field guidelines for remote sensing species-rich grasslands need to be elaborated.

Keywords: *remote sensing; species-rich grasslands; habitat mapping; spectral variation hypothesis; biodiversity; community traits*

5.1 Introduction

Grasslands provide a range of benefits, such as sequestrating carbon, forage production, providing habitat for pollinators and other declining species, nutrient recycling, and flood mitigation, yet they are some of the least protected ecosystems (<5% fall within a protected area), and almost 50% of the world's grasslands are considered degraded (Carbutt *et al.*, 2017). Species-rich grasslands (SRGs) are amongst the most impacted because of conversion to agriculture, with grasslands having larger conversion rates than those of forests (Scholtz and Twidwell, 2022). For example, in the UK, only 3% of initial SRGs remain intact (Plantlife, 2018). While the importance of grasslands is being increasingly recognised by both local small-scale monitoring and restoration initiatives e.g., the Magnificent Meadows project in the UK (Plantlife, 2018) and international organisations, such as the WWF's 'Global Grassland & Savannah Dialogue Platform' (Bardgett *et al.*, 2021), greater emphasis is needed for policy implementation to look towards the conservation of these ecosystems, rather than solely rely on policies surrounding reforestation, for example, for climate change mitigation (Buisson *et al.*, 2021).

New targets tend to focus on the restoration of SRG habitats by recovering sites of previously known grassland locations (Török *et al.*, 2021; Wilsey, 2021). Although this is important, locating and assessing grasslands, among other priority habitats, must be the starting point (as outlined by the EU's biodiversity Strategy, developed from the Aichi Targets) (EEA, 2016). There may be pockets of undisturbed grasslands that can provide resources, such as vital seed banks or refuge areas that organisations and conservationists are unaware of. Therefore, accurate mapping of the distribution and condition of global grasslands is essential to support, maintain, and increase current grassland ecosystems (Buisson *et al.*, 2021). However, mapping projects vary largely by habitat type and country. When considering grasslands, these intricate habitats are often constricted into one general category of "grassland" (for example, acid, neutral and calcareous grasslands are grouped together as one in the Living England Project (Kilcoyne *et al.*, 2022)) or split broadly into "acid", "neutral", and "calcareous" (such as the case in the UK CEH land cover maps (Rowland *et al.*, 2020)) with little consideration between species-rich and species-poor variants, despite providing hugely different ecosystem services. Issues surrounding mapping of biodiverse grasslands could also be linked to the unclear and varying definitions and classification systems that exist for the wider habitat (Dabrowska-Zielinska *et al.*, 2019).

5.1.1 Remote Sensing of Grasslands

Various studies have shown improvements in grassland monitoring and mapping with the inclusion of remote sensing (RS), taking advantage of increased spatial and spectral resolution of airborne and spaceborne sensors (e.g., Andreatta *et al.*, 2022; Rossi *et al.*, 2019; Wachendorf *et al.*, 2018; Wang *et al.*, 2022). As grasslands are mosaiced and transitional by nature, and with features existing on a small spatial scale, there have been previous issues of classifying and monitoring these landscapes with RS (Reinermann *et al.,* 2020; Rossi *et al.,* 2022). Although earlier studies have involved classifications on a broad ecosystem scale (e.g., differentiating grasslands from forests), including the Global Land Cover 2000 (GLC, 2003) and the MODIS global land cover (Friedl *et al.,* 2002) datasets, more nuanced discrimination and classification of intra-grassland classes by RS are currently not wide-scale, despite continued calls to do so (Giri *et al.,* 2005; Raab *et al.,* 2021).

Previously, RS of grasslands has largely used MODIS or Landsat satellites for this monitoring, but research has shown that both high spatial, spectral, and temporal resolution are needed for improved grassland observation (Andreatta *et al.,* 2022; Gholizadeh *et al.*, 2020; Rossi *et al.,* 2022). The use of Sentinel-2 satellites has had more attention since their launch in 2015 and 2017, due to their higher spatial resolution (10 m), spectral resolution (13 bands) and revisit frequency (5 days). These applications of Sentinel-2 satellites can be seen in more recent studies assessing functional, spectral, and species diversity as well as productivity in grasslands (Muro *et al.,* 2022; Rossi *et al.,* 2020; Rossi *et al.,* 2021). Although the use of Sentinel-2 satellites is predicted to increase in use, its applications in grasslands monitoring are still relatively low (Soubry *et al.*, 2021). Image fusion and data combinations with higher spatial resolution satellites, such as PlanetScope (3 m resolution) or Unoccupied Aerial Vehicles (UAV) (cm level resolution) have been shown to improve accuracy in grassland RS applications (Andreatta *et al.,* 2022; Muro *et al.,* 2021). A greater culture of integration and OS make it possible to combine multiple datasets for a more robust and rounded method of grassland observation, rendering it necessary to investigate the use of various RS devices in applications of grassland monitoring.

The varying phenological, functional, and diversity (collectively community) traits of grassland classes are often difficult to separate. Additionally, global classification systems of these habitats are not aligned, resulting in no one discrimination method being applicable to all (Reinermann *et al.,* 2020). Sullivan *et al.* (2010) noted the gradient between improved (agriculturally enhanced through chemical inputs and heavy grazing and/or cutting) – semi-improved (previously agriculturally enhanced or reduced inputs and grazing/cutting) (Swallow, 2016) grasslands over a decade ago, and called for further classifications to be added, demonstrating how difficult the differentiation between grassland classes can be. Evidence shows that grassland trait retrieval is possible by optical RS, through modelling the relationship between traits and spectral responses (e.g., Homolová *et al.,* 2013; Verrelst *et al.,* 2015). Successfully retrieved traits include biochemical properties (such as water content, chlorophyll, and nutrient composition), and biophysical properties (such as leaf area or biomass) (Li *et al.,* 2018; Zhao *et al.,* 2021*b*; Zhang *et al.,* 2023). Therefore, it is crucial to identify what grassland community traits can be accurately estimated across grassland classes from RS to use these variables as predictors in classification, mapping, and monitoring attempts.

5.1.2 Characteristics of Species-Rich Grasslands for Remote Sensing

SRGs have high floristic diversity of >12-15 species per square meter, and are historically differentiated into acidic, neutral, and calcareous (and marshy) classes by their pH values (JNCC, 2010). The characteristics of each species-rich grassland class have not been extensively defined beyond species presence, number of taxa (species richness), and pH level (Pers.obvs.). As determining the number of taxa is vital in confirming the presence of SRGs, it would be crucial to assess the potential of species richness to be reliably represented using RS methods.

The Spectral Variation Hypothesis theorises that greater species diversity of a habitat leads to increased spectral diversity in response (Palmer *et al.*, 2002). This has been demonstrated in a range of grassland conditions (monocultures, farmland, prairie) but less so for floristically diverse grasslands, especially those located in Europe (Polley *et al.*, 2019; Wang *et al.*, 2018a; Zhao *et al.*, 2021*a*). Issues with the hypothesis tend to be found at landscape level, potentially due to species richness of land cover types being affected by differing confounding factors, however, the metric seems particularly plausible in determining alpha (within community) and beta (between community) diversity (Schmidtlein and Fassnacht, 2017; Rocchini *et al.*, 2021).

The hypothesis has previously been shown to be confounded by spatial scale and proportion of bare soil, for example, which influence the spectral response (Gholizadeh *et al.*, 2018; Rossi *et al.*, 2022). Therefore, the inclusion of sufficient predictor variables, such as vegetation indices (VIs), which may account for confounding factors seems pertinent. Due to the high biological diversity found between and within species-rich grassland classes, spectral diversity could have the potential to differentiate grassland classes in future prediction mapping; specifically, for locating SRGs, or, at the very least, separating them from their less diverse counterparts.

Plant functional and structural traits distinguish community and grassland productivity – a key factor in differing agricultural grasslands from semi-natural grasslands, such as SRGs (Hetzer *et al.*, 2021). Traits, such as sward height and above ground biomass (AGB) can often indicate management practices. For example, one way unimproved grasslands are differentiated from improved grasslands is by their grazing/cutting frequency, which can be suggestive of the grassland class (Vickery *et al.*, 2001). Different management practices and species presence across classes will result in these traits varying by grassland type (Magnificent Meadows, 2019). Whilst within SRG classes, acidic and calcareous grasslands are often found to have shorter swards than their neutral counterparts (pers.comms., 2023). Chlorophyll content may also be crucial in determining SRGs. Heavy fertiliser applications on improved grasslands create an artificial greenness of the sward, whilst being cut more regularly reduces dead material, influencing the overall chlorophyll content of the habitat. Chlorophyll content is also vital for detecting temporal differences, a factor that is shown to limit grassland RS, where harvesting agricultural grasslands also alters the chlorophyll content in comparison to natural swards (Möckel *et al.*, 2014; Shellswell, 2017).

5.1.3 Aims and Objectives

The thesis RQ4 was adapted for this chapter, specific to the identified case-study, resulting in the overall research question: Is currently available open-source remote sensing data able to accurately monitor species-rich grasslands and their vulnerable species?

Very little research has categorised a wide range of environmental characteristics across SRGs of varied type and location and therefore assessing certain predictor variables may help in their identification on wider scales. This research aimed to investigate the relationship between species diversity, spectral diversity, and the retrieval success of grassland community traits in SRGs found across Scotland. It is expected that this information could be utilised to improve future prediction and mapping attempts of SRG distributions. Assessments used Sentinel-2, Planetscope, and Micasense MX Red-edge Dual Camera sensors with varying spectral (8 – 13 bands) and spatial scales (8 cm – 10 m).

As such, this chapter aimed to answer the questions:

i) Is there a positive relationship between species diversity and spectral diversity across SRG sites in Scotland?

ii) Can plant community structural and biochemical traits (SPAD-measured chlorophyll-proxy, sward height, fresh weight, and above ground biomass) be estimated using RS across SRG sites in Scotland?

5.2 Methods

5.2.1 Overall Approach

To answer both questions, environmental and remote sensing data collection needed to occur at SRG sites across Scotland. This data that was collected in 2022 came from a reduced number of sites from the previous data collection period in 2021 (described in chapter 4). Only 11 of the original 16 sites were revisited in 2022, due both to time restrictions and some sites not having wholly suitable conditions. For example, at Glensaugh Lower it proved to be very difficult to find a grassland area to survey, with other sites of its classification being more representative of a SRG. Satellite imagery and environmental data was collected at all 11 sites, whereas UAV data was collected only where permissions were granted. The data was then sorted into analysis for investigating the relationship between spectral and species diversity and data that would be used to investigate plant community trait estimation. Although sperate analyses were conducted to answer each question, calculated spectral diversity metrics associated with the species-spectral diversity relationship were investigated in the trait estimation analysis to utilise the availability of a large dataset. An overview of the methodological approach can be seen in Figure 5-1.

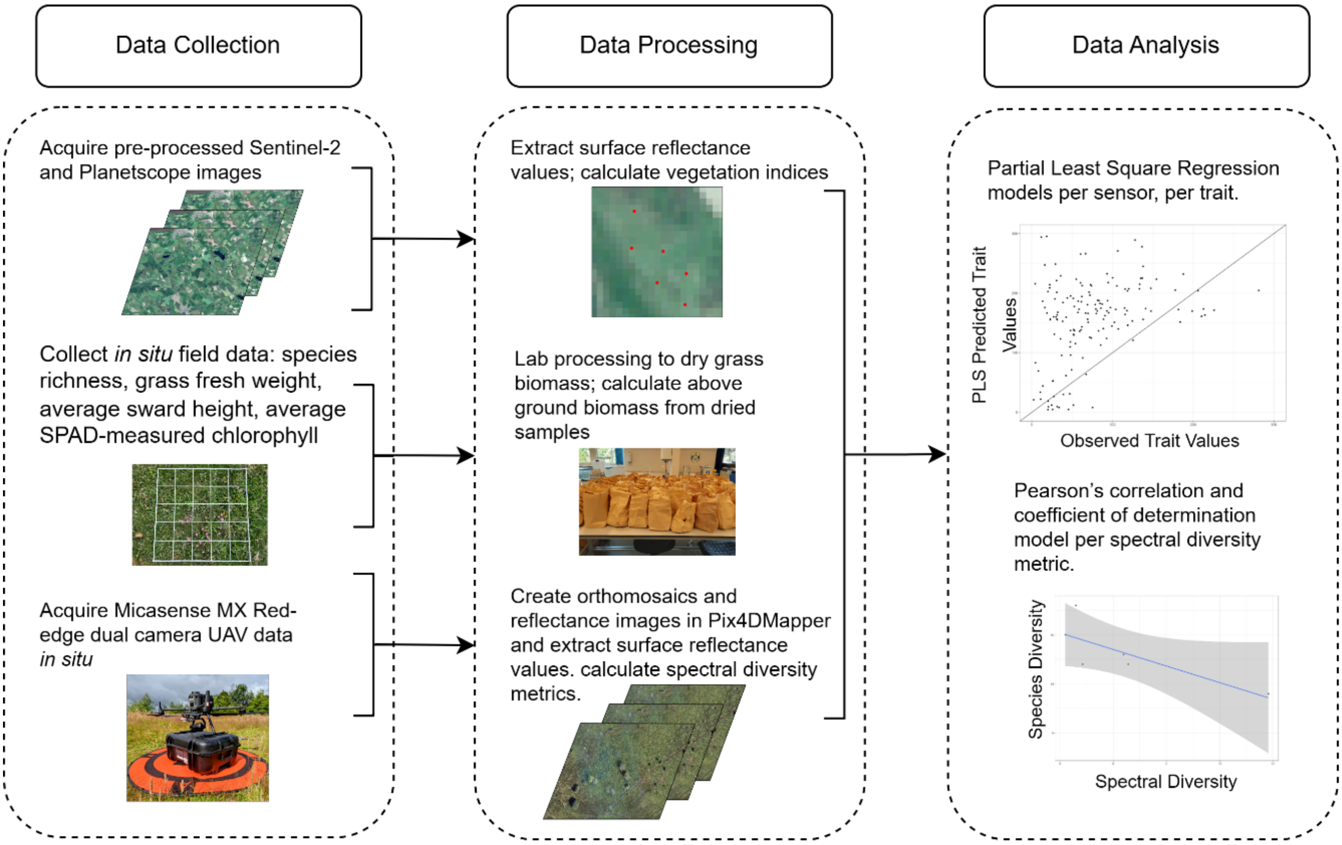


Figure 5-1. Processing steps involved in model creation for predicting species richness and grassland community traits in species-rich grasslands across Scotland.

5.2.2 Study Sites

Over the grass growing season (May – August) in 2022, eleven SRGs were sampled three times across Scotland (Figure 5-2). Sites were chosen across the country to incorporate a range of SRG classes and variable climatic conditions that affect species presence and diversity. Open-source data (from Butterfly Conservation, Nature Scot, and the UK Centre of Ecology and Hydrology) on previous and potential SRG grassland locations were used to identify sites.

The sites varied in their species richness and community traits, including both wet and dry, and semi-improved versus more unimproved sites. This allowed greater representation of semi-natural grasslands across Scotland. Satellite data was captured across neutral, calcareous, acid, and coastal SRGs, whereas UAV data was only captured across seven neutral SRGs; limited by access, permissions, and weather (Table 5-1). The same data was also collected for these eleven sites, plus an additional five, over 2021, however, was not used in the analysis due to inconsistencies in site coverage.

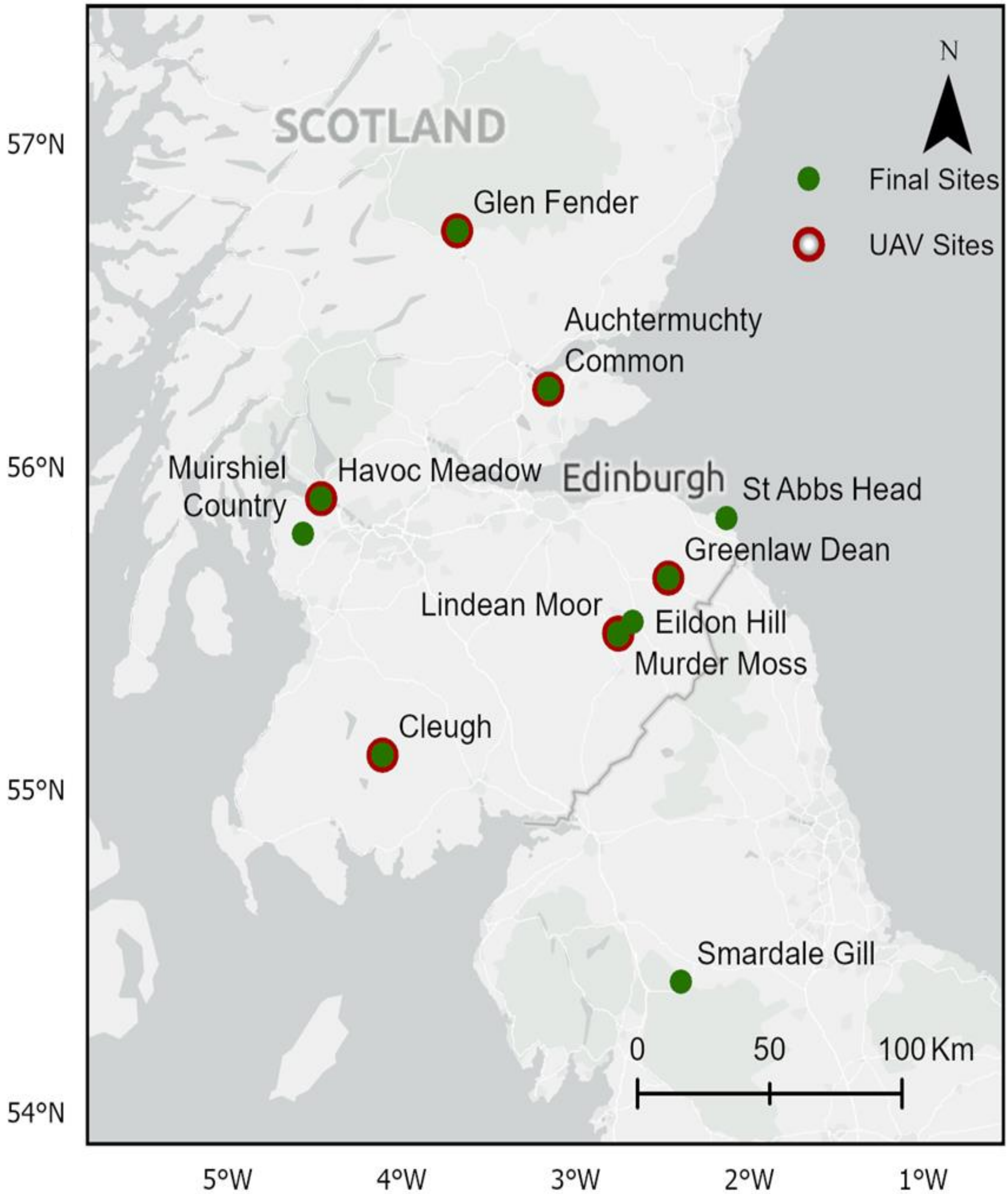


Figure 5-2. Species-rich grassland sites located across Scotland that were sampled in 2022. UAV data was acquired at seven of the eleven sites (highlighted in red). Lindean Moor and Murder Moss sites are adjacent to one another and represented by a single point.

Table 5-1. Field sites where remote sensing, community trait, and species richness data was collected in 2022. UAV (highlighted in bold) data for the three campaigns across 2022, with corresponding Sentinel-2 (S2) and Planetscope (PS) acquisition dates per site. See Appendix D-1 for corresponding site code names.

Site	MAY			JUNE/JULY			AUGUST		
	Field Survey	*S2*	*PS*	*Field Survey*	*S2*	*PS*	*Field Survey*	*S2*	*PS*
AC	**24/05/22**	19/05/22	08/05/22	**12/07/22**	18/07/22	07/07/22	**16/08/22**	10/08/22	11/08/22
CL	**19/05/22**	27/05/22	26/05/22	**30/06/22**	04/07/22	10/07/22	**04/08/22**	10/08/22	04/08/22
EH	17/05/22	24/05/22	27/05/22	28/06/22	18/07/22	07/07/22	02/08/22	10/08/22	09/08/22
GD	**17/05/22**	14/05/22	22/06/22	**28/06/22**	18/07/22	07/07/22	**02/08/22**	10/08/22	10/08/22
GF	25/05/22	25/04/22	05/06/22	**13/07/22**	09/07/22	10/07/22	**17/08/22**	10/08/22	17/08/22
HM	26/05/22	04/06/22	29/05/22	**14/07/22**	09/07/22	10/07/22	**18/08/22**	10/08/22	11/08/22
LM	**16/05/22**	24/05/22	08/05/22	**27/06/22**	18/07/22	22/06/22	**01/08/22**	10/08/22	09/08/22
MM	**16/05/22**	24/05/22	08/05/22	**27/06/22**	18/07/22	22/06/22	**01/08/22**	10/08/22	09/08/22
MP	26/05/22	04/06/22	29/05/22	14/07/22	09/07/22	18/07/22	18/08/22	20/08/22	12/08/22
SA	23/05/22	24/05/22	24/05/22	11/07/22	18/07/22	10/07/22	15/08/22	10/08/22	10/08/22
SM	18/05/22	14/05/22	18/05/22	05/07/22	16/07/22	09/07/22	03/08/22	10/08/22	10/08/22

5.2.3 *In Situ* Data Collection and Lab Processing

At each site, the transect origin point was determined using a random sampling approach, decided by randomly throwing a quadrat. A 250 m W transect was used (Figure 5-3a) to gain representative community trait variation over the habitat (Byrne *et al.*, 2018; Milner and Sharp, 2014). The transect length was constrained by varying site sizes (found between <1 – 326 ha) and surveyable areas (i.e., areas devoid of trees, shrub, site boundaries, and water features), so 250 m was chosen to ensure a final consistent transect length between sites. A total of six 0.5 m^2 quadrats were used along the transect at each 50 m section to measure the community traits and count the species richness. Overall, 198 quadrats were surveyed across the range of grassland classes. However, due to missing information, the dataset is made up of 195 quadrats (Figure 5-4).

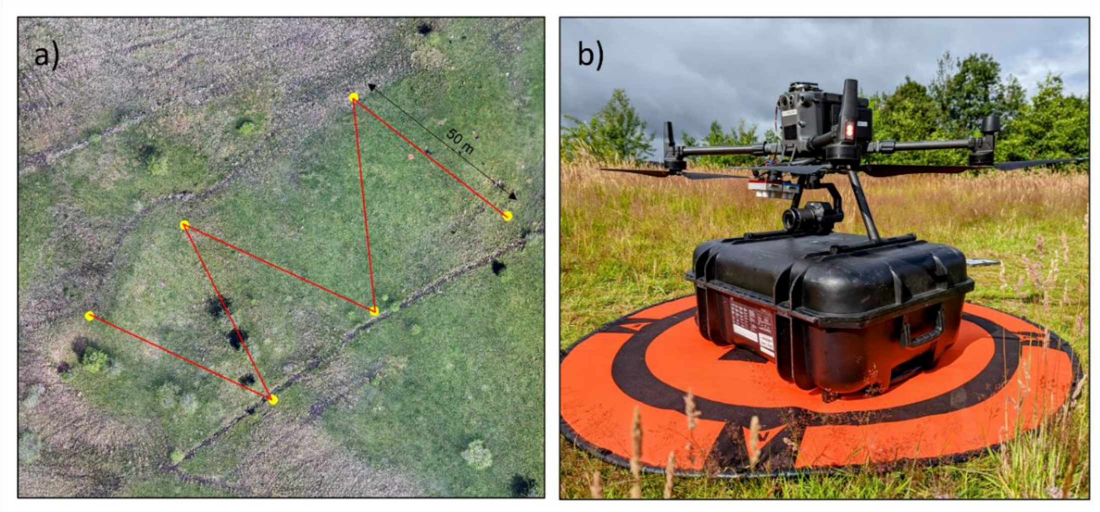


Figure 5-3. A) Orthophoto of a site detailing the W transect methodology employed to survey each site. The W transects were made up of 5 x 50 m (total = 250 m) splits, each ending with a 0.5 m2 quadrat; b) DJI Matrice 300 RTK UAV platform that was used for all acquisitions with a Micasense Rededge MX 5.5 dual camera and DJI Zenmuse P1 sensor.

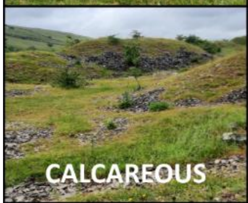
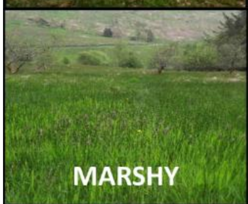
Species-rich Grassland Class	Description and Indicator Species
 ACID	• Found on free draining soils • Soil pH of <5.5 • May have lower floral diversity • Indicator species include fine leaved grasses (e.g. matt grass, sheep's fescue), common tormentil, heath bedstraw
 CALCAREOUS	• Rare in Scotland • Found on thin, calcium soils • Soil pH of >7.5 • Indicator species include blue moor grass, quaking grass, common rock-rose, wild thyme
 COASTAL	• Not a previously defined class (usually associated with acidic soils, not to be confused with dune habitats) • Distinct in floral diversity due to coastal air • Presence of coastal species, such as sea thrifts and sea campions
 MARSHY	• Found on poorer draining soils • Seasonally wet with varying water table height • Indicator species include rushes and sedges, purple moor grass, meadow sweet, marsh marigold, marsh thistle
 NEUTRAL	• Most associated with typical wildflower meadows • Soil pH usually 5.5 – 7.5 • Indicator species include crested dogs tail, meadow fox tail, cock's foot, yarrow, bird's foot trefoil

Figure 5-4. Varying community and species compositions across five types of species-rich grasslands in Scotland.

Recorded variables were selected based on the community traits (above ground biomass, sward height, fresh weight, and SPAD-measured chlorophyll-proxy) that vary between SRG classes. At each subsite, species richness was determined by counting the total number of species within a quadrat (Rossi *et al.*, 2022). Sward height was recorded by taking five random measurements of the tallest vegetation touching a tape measure before the removal of all above ground plant material in a quadrat to grazing level (recorded as fresh weight) for the subsequent determination of above ground biomass (AGB). On return to the laboratory, this material was oven dried (oven LTE OP250 40°C to 250°C) at 70° C until a constant weight was reached. The dry weight was then calculated per area unit. Chlorophyll content could not be measured with a spectrophotometer due to equipment and time limitations. As such, a chlorophyll proxy was utilised, measured with a SPAD (Soil Plant Analysis Development) meter (Konica Minolta SPAD-502 Plus) on one leaf of five randomly chosen individual grasses, *in situ* within the quadrat. SPAD measurements were only taken from grass species as these were consistently the dominant plant groups within a quadrat. The specific leaf of each individual plant was chosen as the first leaf from the ground that would entirely cover the SPAD sensor. The SPAD meter was calibrated before each use. The five values were averaged per plot to give a SPAD-measured chlorophyll-proxy (SMCP) value.

5.2.4 Satellite and Unoccupied Aerial Vehicle Acquisitions

We used sun-synchronous Sentinel-2A (orbital altitude 786 m; orbital inclination 98.62°) and Planetscope (orbital altitude 475 – 525 km; orbital inclination 98°) satellites. Sentinel-2A (S2) data was acquired due to its high spectral resolution of 13 bands and medium spatial resolution of 10 – 60 m, as well as being open-source data. Level-2a atmospherically corrected surface reflectance imagery was acquired using Google Earth Engine for each site, as close as possible to survey dates (+/- 0 – 31 days, see Table 5-1 for corresponding dates). Specific S2 bands (B2, B3, B4, B8 at 10 m resolution and B5, B6, B7, B8A, B11, B12 at 20 m resolution) that are associated with vegetation characteristics were selected for use in image classification and resampled (nearest neighbour) to the highest 10 m resolution. Planetscope (PS) surface reflectance data were also acquired due to its higher spatial resolution of 3 m; however, it has a reduced spectral resolution of eight bands. This allowed comparison between spectral and spatial resolution limitations. The raster package (Hijmans, 2023) in R (v 3.6.3) was used to extract the satellites' reflectance values for each pixel corresponding to the quadrat location.

A DJI Matrice 300 RTK UAV platform (Figure 5-3b) was used and flown at 50 m, with an 80% front and side overlap, to take hyperspatial (8 cm), multispectral (10 band) acquisitions with a Micasense Rededge MX 5.5 dual camera. The use of 7 – 10 ground control points were installed on each site for accurate positioning, using a Leica Viva GS08 GNSS receiver. The reflectance was calibrated in field using a white reflectance panel (RP06-2102083-OB) to incorporate the day's lighting conditions. This sensor allows bands to be synergised with Sentinel-2 bands, as well as providing hyperspatial resolution multispectral data. Pix4D Mapper (v 4.6.4) was used to create orthomosaics and surface reflectance images of the sites (georeferenced between 0.009 – 0.054 m). ArcGIS Pro (v 3.0.36057) was used to create 0.5 m^2 polygons over the quadrats. The R raster package (Hijmans, 2023) was then used to extract the pixel reflectance values per quadrat and average them across the polygon areas for each of the 10 bands.

5.2.5 Vegetation Indices

Vegetation indices (VIs) were calculated from the spectral reflectance values, as often the ratio between bands can be more sensitive to variables on the ground (as seen in the positive relationship between the Normalised Difference Vegetation Index (NDVI) and AGB) (Meng et al., 2017; Zhang et al., 2015). As such, several VIs were calculated: NDVI, including the red-edge NDVIs and indices, which represent information on chlorophyll content, such as Sentinel-2 Red-edge Position Index (S2REP); the Enhanced Vegetation Index (EVI); and the Greenness Vegetation Index (GVI) (Table 5-2).

Table 5-2. Vegetation indices calculations per sensor. Band numbers listed by central wavelength (nm).

Vegetation Indices	Sentinel-2A	PlanetScope	Micasense MX Red-Edge Dual camera	Source
NDVI	$(B8_{(842)}-B4_{(665)})/(B8_{(842)}+B4_{(665)})$	$(B8_{(865)}-B6_{(665)})/(B8_{(865)}+B6_{(665)})$	$(B4_{(842)}-B3_{(668)})/(B4_{(842)}+B3_{(668)})$	Imran *et al.*, 2020 ; Peciña et al., 2021 ; Qin *et al.*, 2021
EVI	$2.5(B8_{(842)}-B4_{(665)})/(B8_{(842)}+6\ B4_{(665)}-7B2_{(490)}+1)$	$2.5(B8_{(865)}-B6_{(665)})/(B8_{(865)}+6\ B6_{(665)}-7B2_{(490)}+1)$	$2.5(B4_{(842)}-B3_{(668)})/(B4_{(842)}+6B3_{(668)}-7B1_{(475)}+1)$	Peciña et al., 2021; Qin *et al.*, 2021; Zou *et al.*, 2022
GVI	$(B8_{(842)}-B3_{(560)})/(B8_{(842)}+B3_{(560)})$	$(B8_{(865)}-B4_{(565)})/(B8_{(865)}+B4_{(565)})$	$(B4_{(842)}-B2_{(560)})/(B4_{(842)}+B2_{(560)})$	Peciña et al., 2021
S2REP	$705+35((2(B7_{(783)}+B4_{(665)})-(B5_{(705)})/(B6_{(740)}-B5_{(705)}))$			Li *et al.*, 2021; Zou *et al.*, 2022
$NDVI_{Red\text{-}Edge\ 1}$	$(B8_{(842)}-B5_{(705)})/(B8_{(842)}+B5_{(705)})$	$(B8_{(865)}-B7_{(705)})/(B8_{(865)}+B7_{(705)})$	$(B4_{(842)}-B9_{(705)})/(B4_{(842)}+B9_{(705)})$	Li *et al.*, 2021; Peciña et al., 2021
$NDVI_{Red\text{-}Edge\ 2}$	$(B8_{(842)}-B6_{(740)})/(B8_{(842)}+B6_{(740)})$		$(B4_{(842)}-B10_{(740)})/(B4_{(842)}+B10_{(740)})$	Li *et al.*, 2021; Peciña et al., 2021
$NDVI_{Red\text{-}Edge\ 3}$	$(B8_{(842)}-B7_{(783)})/(B8_{(842)}+B7_{(783)})$		$(B4_{(842)}-B5_{(717)})/(B4_{(842)}+B5_{(717)})$	Li *et al.*, 2021; Peciña et al., 2021
$ND_{Red\text{-}Edge1}$	$(B6_{(740)}-B5_{(705)})/(B6_{(740)}+B5_{(705)})$		$(B10_{(740)}-B9_{(705)})/(B10_{(740)}+B9_{(705)})$	Li *et al.*, 2021; Imran *et al.*, 2020
$ND_{Red\text{-}Edge2}$	$(B7_{(783)}-B5_{(705)})/(B7_{(783)}+B5_{(705)})$		$(B5_{(717)}-B9_{(705)})/(B5_{(717)}+B9_{(705)})$	Li *et al.*, 2021; Imran *et al.*, 2020
NDII	$(B8_{(842)}-B11_{(1610)})/(B8_{(842)}+B11_{(1610)})$			Li *et al.*, 2021 ; Qin *et al.*, 2021

5.2.6 Analysis

5.2.6.1 Species – Spectral Diversity Relationship

The Micasense UAV data (*n=114*) was used to evaluate the relationship between spectral and species diversity with common spectral diversity metrics including: the standard deviation (SD) and coefficient of variation (CV). These metrics have often shown to have the most predictive power in diversity estimations (Imran *et al.,* 2021; Peng *et al.,* 2019; Wang *et al.,* 2018a). Other metrics, such as Convex Volume of Hull is more easily influenced by outliers (Tassi *et al.,* 2022). Clustering methods were not used; RaoQ is more applicable for assessing functional (abundance weighted) diversity which could not be measured in this study, whilst the Spectral Species approach is less suitable for cross-site studies (Féret and Asner, 2014; Rocchini *et al.,* 2017; Rossi *et al.,* 2022).

Both SD and CV metrics were calculated by averaging individual SD and CV values across the pixels in a quadrat (0.5 m^2). For this reason, only Micasense data was applicable for this objective as Sentinel-2 and Planetscope pixel areas exceeded the quadrat area (10 m^2 and 3 m^2 versus 0.5 m^2). Extrapolating the satellite data over an increased area was not applicable for these sites due to the mosaiced habitats. An NDVI soil mask (pixel values ≥ 0.4) and a NIR shade mask (pixel values ≥ 0.22) were applied to the UAV data (Schweiger and Laliberté, 2022). However, it was observed that most quadrats were fully vegetated (>99% of pixels), so soil was unlikely to be an influencing factor. Due to the structure and composition of the grassland species, and the avoidance of trees/shrub close to quadrats, shadows were also not likely to confound the reflectance values (>99% of pixels were above threshold). The average SD and CV of spectral reflectance values across the visible and near infrared wavelengths (444–842 nm) were used. CV was calculated as the standard deviation/mean reflectance value at a specific band, then averaged across all relevant bands per sensor, for each quadrat (as in Eq. 1).

$$CV_{quadrat} = \Sigma(SD_{wavelength}/Mean_{wavelength})/\text{number of bands} \quad (1)$$

Species diversity was measured as species richness, a commonly used species diversity metric (Peng *et al.,* 2019). The relationships between spectral – species diversity was tested using a Pearson Correlation coefficient ©. The assumptions for both metrics were visualised and tested. The log transformation of SD was used for Spectral Diversity SD to assure the assumptions were met. The predictive power of spectral diversity on species and trait diversity was then estimated with the coefficient of determination (R^2).

5.2.6.2 Species-Rich Grassland Community Trait Retrieval

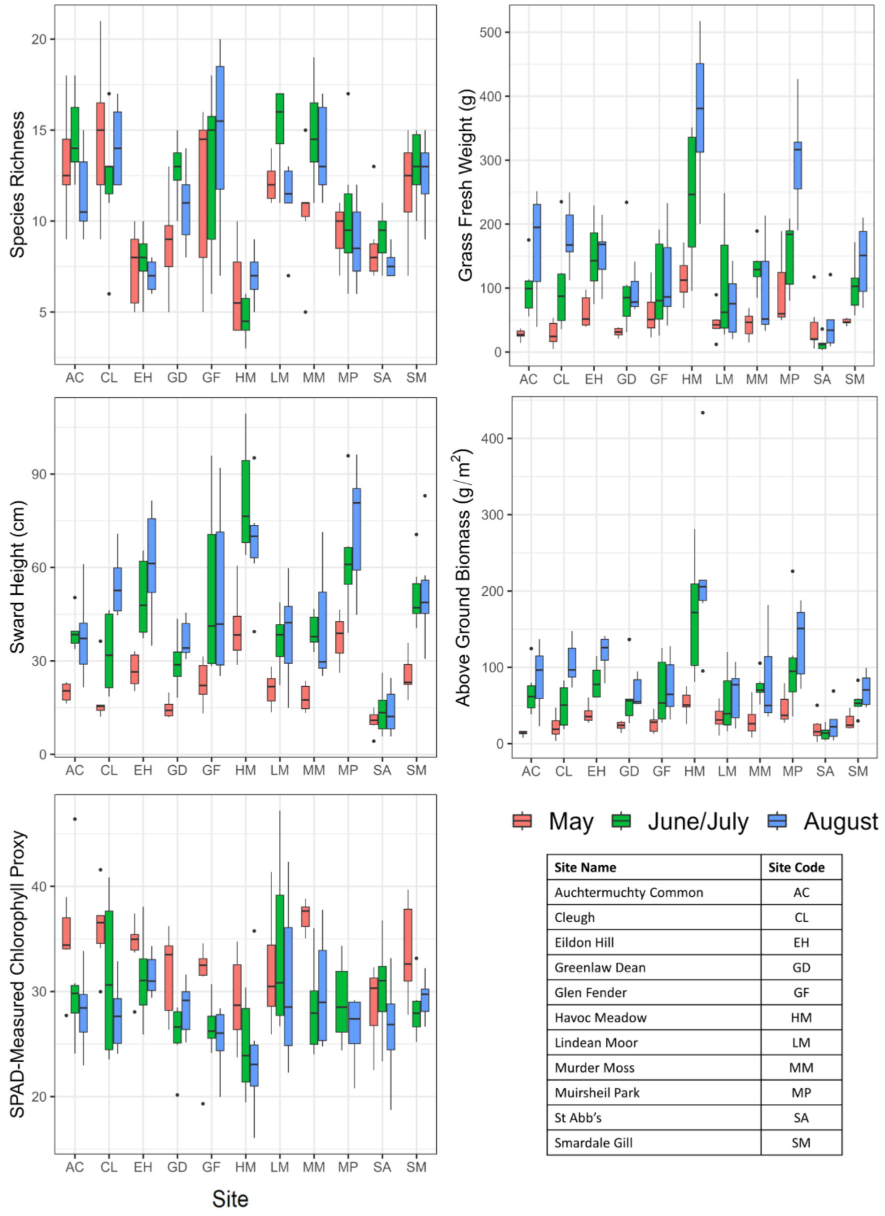
Both satellite (S2 and PS) and UAV (Micasense MX Red-Edge dual camera) data were used to investigate grassland community trait estimation from RS data, to allow for a spatial resolution comparison. The partial least square regression model (PLS) was chosen to evaluate RS data in predicting grassland community structural (AGB, sward height, and fresh weight) and functional (SMCP) traits. Partial least square regression is a non-parametric model which works well with highly colinear data, such as environmental data. This model is commonly used for trait estimation from spectral data and consistently outperforms other models such as principal component analysis in prediction analyses for this type of evaluation (Capolupo *et al.*, 2015; Pang *et al.*, 2020; Schweiger *et al.*, 2017; Zhang *et al.*, 2022).

Both surface reflectance data of relevant bands in vegetation analysis (spectral range) and vegetation indices were used as predictor variables, as seen in previous studies (Capolupo *et al.*, 2015; Pang *et al.*, 2020; Zhao *et al.*, 2021*b*). Spectral diversity metrics of the MS were also included in the MS models to see if this improved trait predictions, as a unique method of trait retrieval. It was not possible to compute spectral diversity metrics for S2 and PS data per quadrat. Although these metrics could have been calculated for S2 and PS data per site or class, this was not done as the data would average to eleven or five (respectively) data points, which was not suitable for train/test data partitioning. PLS models were created for S2 data (PLS-S2), PS data (PLS-PS), and for Micasense data (PLS-MS). For each dependent variable, the model was originally created with all predictor components. Each sensor had differences in the number of bands, VIs, and whether there were spectral metrics or not. As such, multiple combinations of the predictor components were tested in each model to investigate the most relevant bands and VIs for the highest prediction accuracy per trait estimation. For example, model iterations included the sole removal of NIR bands, the sole removal of RGB bands, the sole removal of SWIR bands, and the sole removal of edge bands (such as the red-edge bands). This is because different regions of the spectrum hold varying information on plant properties depending on the interaction at particular wavelengths (Ge *et al.*, 2019). All models were trained and tested with a 70:30 percentage split. A 10-fold cross validation was performed on each model and the best tuned model was chosen as the final model per trait prediction.

5.3 Results

5.3.1 Site Descriptive Statistics

SRG species numbers and community trait values varied largely across the SRG sites due to the heterogeneity seen both within and between classes (Figure 5-5). Species richness ranged from 3 to 21 species found in a quadrat, across the seven neutral grassland sites. Average sward height of a quadrat ranged from 4.24 cm to 109.38 cm. Total amount of AGB per plot ranged from 2.16 g/m^2 to 433.6 g/m^2 of a quadrat. Average SMCP ranged from 16.06 to 47.2 in a quadrat. Total grass fresh weight per plot ranged from 3.1 g to 517.5 g.



Site Name	Site Code
Auchtermuchty Common	AC
Cleugh	CL
Eildon Hill	EH
Greenlaw Dean	GD
Glen Fender	GF
Havoc Meadow	HM
Lindean Moor	LM
Murder Moss	MM
Muirsheil Park	MP
St Abb's	SA
Smardale Gill	SM

Figure 5-5. Variation in grassland community traits (species richness (number of species), above ground biomass (g/m^2), average chlorophyll, grass fresh weight (g), average sward height (cm), and SPAD-measured chlorophyll proxy across 11 species-rich grassland sites in Scotland.

5.3.2 Spectral-Species Diversity Relationship

The relationship between species richness and spectral diversity was measured for seven of the 11 SRG sites. The Pearson's correlation showed that there was no significant relationship between species diversity and spectral diversity SD ($t = -0.119$, $df = 112$, $p = 0.211$), and spectral diversity SD showed a low predictive power for species diversity ($R^2 = 0.0141$). There was no significant relationship between species diversity and spectral diversity CV ($t = 0.139$, $df = 112$, $p = 0.141$), and spectral diversity CV showed a low predictive power for species diversity ($R^2 = 0.0192$).

5.3.3 Estimations of Grassland Community Functional and Structural Traits

The PLS models showed mostly low predictive power across the remote sensing devices per trait. The Micasense data had the strongest predictive power for all traits: sward height (R^2 = 0.545, RMSE = 13.56 cm, removing the RGB and SWIR surface reflectance band values and spectral diversity metrics from the model), AGB (R^2 = 0.221, RMSE = 48.26 g/m^2, removing the RGB and SWIR surface reflectance bands from the model and the red-edge VIs), fresh weight (R^2 = 0.235, RMSE = 98.74 g, removing the RGB and SWIR band values from the model), and SMCP (R^2 = 0.167, RMSE = 5.10, removing only the blue band value from the model). Planetscope consistently performed poorly for trait prediction (Table 5-3; Figure 5-6).

Table 5-3. Species-rich grassland community trait estimation using spectral reflectance values, vegetation indices, and spectral diversity (for Micasense) as predictor variables from the best models per trait of remote sensing sensors: Sentinel-2, Planetscope, and Micasense MX Red-edge dual camera. Final models for best trait estimation are in bold.

Traits	**Sentinel-2**			**PlanetScope**			**Micasense**		
	N	R^2	RMSE	*n*	R^2	RMSE	*N*	R^2	RMSE
Sward Height (cm)	195	0.201	19.92	195	0.150	20.98	114	**0.545**	**13.56**
Above Ground Biomass (g/m^2)	195	0.199	49.83	195	0.159	56.09	114	**0.221**	**48.26**
Fresh Weight (g)	194	0.137	106.31	194	0.097	76.06	114	**0.235**	**98.74**
SPAD-measured Chlorophyll-proxy	189	0.062	5.48	189	0.021	5.99	114	**0.167**	**5.10**

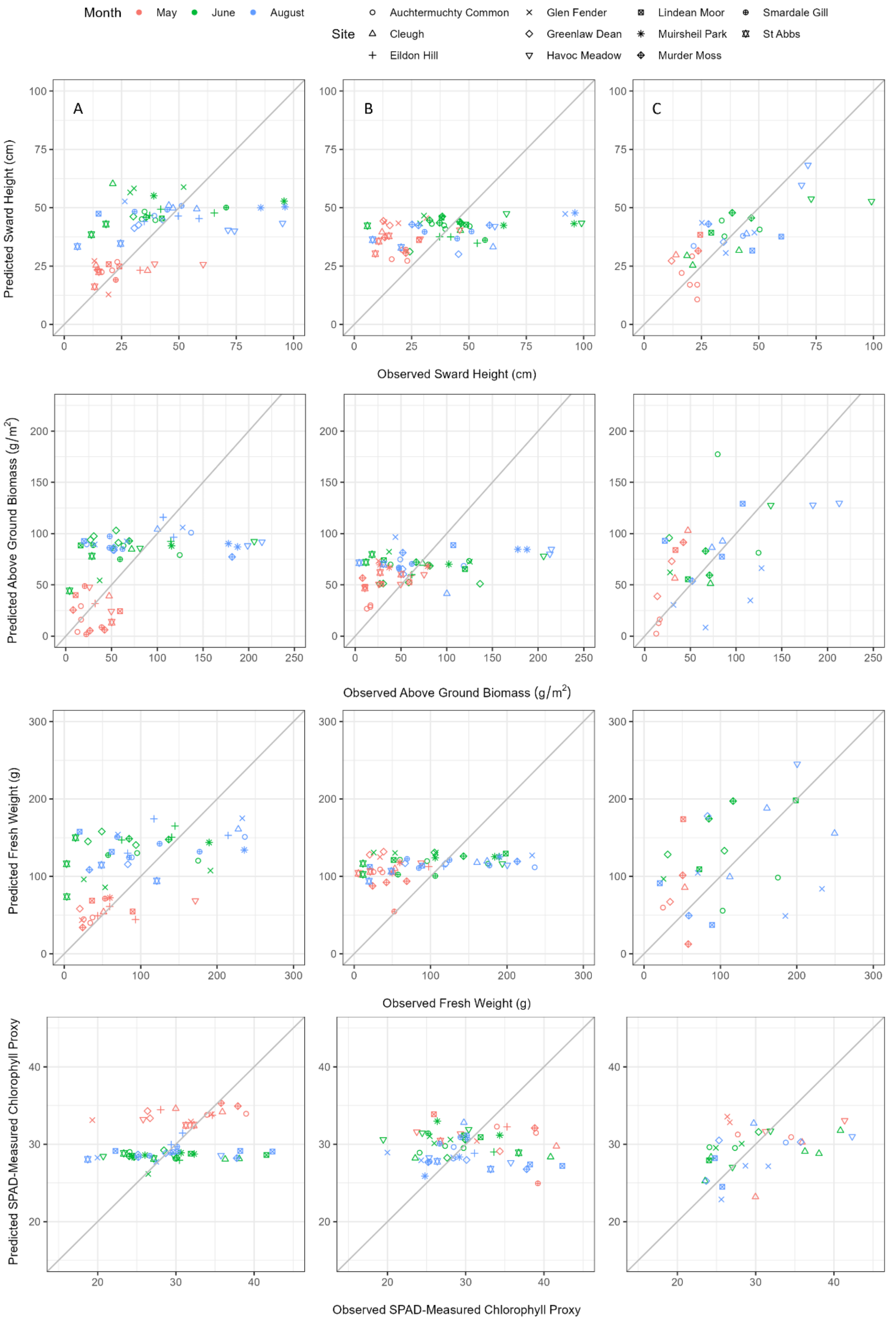


Figure 5-6. Predicted versus observed sward height (cm), above ground biomass (g/m^2), fresh weight (g), and chlorophyll content (SPAD measured) from (A) Sentinel-2 in 10 m resolution, (B) Planetscope in 3 m resolution, and (C) Micasense MX Red-Edge Dual Camera in 8 cm resolution models.

The addition of the MS spectral diversity metrics into the PLS-MS models did slightly improve prediction estimates for AGB, fresh weight, and SMCP but reduced the prediction power of the model for sward height (Table 5-4). The final models for trait prediction are: PLS-MS for fresh weight, AGB, sward height, and SMCP (Table 5-5).

Table 5-4. Partial least square regression model results for grassland trait estimation with and without the inclusion of spectral diversity metrics derived from a Micasense MX red-edge dual camera.

Trait	With spectral diversity metrics		Without spectral diversity metrics	
	R^2	*RMSE*	*R^2*	*RMSE*
Sward Height (cm)	0.305	16.54	0.545	13.56
Above Ground Biomass (g/m^2)	0.221	48.26	0.130	47.74
Fresh Weight (g)	0.235	98.74	0.209	99.07
SPAD-measured Chlorophyll-proxy	0.167	5.10	0.160	5.13

Table 5-5. Final models of species-rich grassland community trait estimation with corresponding best sensor for prediction and the combination of final inputted predictor variables.

Community Trait	Final Sensor for Model	Inputted Predictor Variables		
		Spectral Bands	*Vegetation Indices*	*Spectral Diversity Metric*
Sward Height (cm)	Micasense (PLS-MS)	NIR, Red-Edges	NDVI, GVI, EVI, NDVIRE1, NDVIRE2, NDVIRE3, NDRE1, NDRE2	None
Above Ground Biomass (g/m^2)	Micasense (PLS-MS)	NIR	NDVI, GVI, EVI, NDVIRE1, NDVIRE2, NDVIRE3	CV and SD
Fresh Weight (g)	Micasense (PLS-MS)	NIR, Red-Edges	NDVI, GVI, EVI, NDVIRE1, NDVIRE2, NDVIRE3, NDRE1, NDRE2	CV and SD
SPAD-measured Chlorophyll-proxy	Micasense (PLS-MS)	Blue-444, Red-650, Green-531, Red-Edge 3, NIR, Green, Red	NDVI, GVI, EVI, NDVIRE1, NDVIRE2, NDVIRE3	CV and SD

5.4 Discussion

This study aimed to investigate the retrieval of species richness and other grassland community plant traits by RS with the intention of improving future monitoring and mapping of SRGs. It is one of few extensive explorations of RS data, across spectral and spatial scales, in a broad range of entirely semi-natural SRG grassland sites in Scotland.

5.4.1 Species-Spectral Diversity Relationship

Our results showed that there was no significant relationship between species richness and spectral diversity across the seven study sites where this was investigated. Both spectral diversity metrics (CV and SD) were not good predictors of species richness and explained less than 2% of variation seen. This challenges the Spectral Variation Hypothesis. Although there has been a range of studies (Imran *et al.*

2021; Wang *et al.,* 2018a; Zhao *et al., 2021a)* investigating the Spectral Variation Hypothesis, none have been so extensive across a range of semi-natural heterogenous grasslands, at such a fine spatial scale as this (Rossi *et al.,* 2022).

Our research aligns with many similar studies in that there is no clear relationship between spectral and species diversity in grasslands (e.g., Thornley *et al.*, 2023). Conti *et al.* (2021) found negative relationships between their spectral diversity and species diversity metrics, supporting the negative relationship between spectral diversity SD and species richness found in our study (although this was not significant). Where studies have demonstrated positive relationships between spectral and species diversity, these are found in more homogenous, species-poor, or experimental grassland plots (e.g., Wang *et al.,* 2018a; Zhao *et al.,* 2021*a*). Imran *et al.* (2021) summarised these differences well with their study highlighting the reduced prediction power of spectral diversity metrics in a species-rich grassland versus a species-poor grassland.

Many confounding factors could result in the low confidences found in this relationship, including biomass, species abundances, sward structure, bare soil presence, and spatial scale issues (Gholizadeh *et al.,* 2018; Rocchini *et al.,* 2014). However, when originally choosing modelling parameters to investigate the spectral variation hypothesis, biomass, sward structure, and sward height variation (SD) were included in a linear-mixed affect model and found to have little influence on variation (<5%). As such, these variables were not included as confounding factors in a linear model investigating the relationship between species and spectral diversity. Both spatial and spectral scale are known limitations in grassland RS and there is potential that even multispectral sensors, such as Sentinel-2 may not provide enough information across bands, as reflectance values are given as the value per band's central wavelength. Consequently, for this study our spectral diversity metrics were calculated across 10 bands (multispectral) compared to studies, such as that by Möckel *et al.* (2016), where metrics were calculated across 245 bands (hyperspectral). As such, the reduced amount of surface reflectance (spectral diversity) information across wavelengths may influence the results seen in our study. Spatial scale may have impacted the variety of results across studies too, as it appears to affect parts of the spectrum differently. For example, increased spatial scale resulted in a more positive relationship between spectral and species diversity across the VIS region compared to a weaker relationship across the NIR region (Imran *et al.,* 2021).

A meta-analysis by Thornley *et al.* (2023) identified and summarised the inconsistencies found in the literature. At canopy level, grasslands are prone to quick responses to environmental and management change. Our study sites did consist of a range of management regimes: seven were lightly grazed, two included cutting regimes (cut one year and grazed/left the next), two were unmanaged, and one had the use of herbicide application. One site was accidentally trampled in the late summer.

Phenological stage is also key in grassland monitoring and may have influenced the results, as we measured across a season (Figure 5-7). Thornley *et al.* (2022) found that phenological diversity (defined as the number of phenological stages in a plot) confounded the spectral variation hypothesis, suggesting that the timing of data collection be crucial in determining whether a relationship between spectral and species diversity can be seen. Species type (e.g., graminoids, forbs, legumes, and bryophytes) will also affect the spectral response of a grassland community, as well as shorter responses to extreme weather events (drought/flooding), and diseases (Fassnacht *et al.,* 2022).

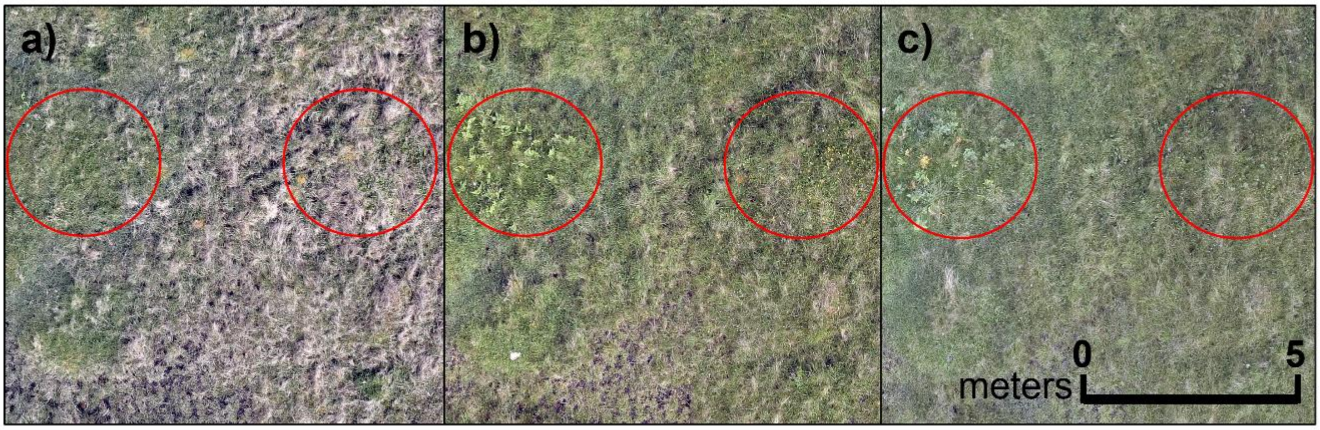


Figure 5-7. Phenological differences at one site at a) May, the start of surveying, b) end of June, the peak of the survey season, and c) August, the end of surveying season, including new growth, increased flowering (circled in red), green-up, then dye-off (yellowing) of vegetation. Images captured with a Zenmuse P1 camera flown at 50 m.

The relative abundance of each species may also confound the species – spectral diversity relationship. Indices that account for species abundance and evenness (Simpson's diversity and Shannon's diversity) have shown positive relationships with spectral diversity (Wang *et al.,* 2018b). Potentially the impact of species abundance is high in heterogenous grasslands, therefore, the dominance of certain species may interfere with the interpretation of the species-spectral diversity relationship.

Other spectral diversity metrics could be tested to see how these influence the results of the diversity estimates, for example clustering techniques, such as the Spectral Species approach. These were not tested here as are less widely applied to the Spectral Variation Hypothesis (only 6/20 studies used a "complex" clustering approach in the meta-analysis by Thornley *et al.* (2023)) and assume that one pixel is one species. We did not think this suitable for our analysis where one pixel, even at 8 cm, was likely made up of more than one species due to sizes of certain herb species and of species overlapping (Rocchini *et al.,* 2022). However, if the approaches can bypass the issues of cross-site comparisons this could be more appropriate for the handling of outliers that would be seen in these heterogenous grasslands.

5.4.2 Grassland Community Trait Estimation

We found low R^2 values (all but one <0.5) from the PLS models for trait estimation, inconsistent with recent studies (Qin *et al.,* 2021; Zhang *et al.,* 2018; Zhao *et al.,* 2021*b*). AGB has often been well predicted by VIs, however, the data collected in these studies comes largely from homogenous agricultural or experimental grasslands to assess productivity or grazing intensity (Jiménez-Jiménez *et* *al.,* 2022; Jin *et al.,* 2014; Zhang *et al.,* 2018; Qin *et al.,* 2021). Evidence also suggests that in more complex vegetation types, with diverse structural and functional traits, the retrieval accuracy (a representation of the similarity between the actual and predicted trait values) of certain properties, such as LAI, and both fresh and dry biomass was reduced (e.g., Imran *et al.,* 2020; Moeckel *et al.,* 2017; Lussem *et al.,* 2019). It appears that the application of predictive models where they have been successful (high R^2 value over 0.5) may not be suitable across various grassland types. This may explain the range in model performances that are seen across studies, including our own low R^2 values, with large variation both within and across sites (Grüner *et al.,* 2019).

Most quadrats had a high proportion of flowering plants, which has previously been shown to negatively affect predictive power of statistical models in the retrieval of plant traits (Schiefer *et al.,* 2021). Zhang *et al.* (2022) found that model prediction for structural properties, such as dry biomass and plant height was similarly low (explaining less than 45% variance) in natural heterogenous grasslands, however, achieved high R^2 values for estimating chlorophyll content. Unlike in our study, the SMCP was most poorly predicted with all models explaining less than 20% of the variation. This could be explained by the number of heterogenous study sites, or the methodology associated with measuring chlorophyll. Spectrophotometers are known chlorophyll extractors in the lab. A SPAD meter measures the amount of light transmitted by certain wavelengths and is proportional to the amount of chlorophyll within a leaf (Süß *et al.,* 2015). As such, it is instead a proxy of chlorophyl content. However, Ludwig *et al.* (2022) demonstrated that after calibration the relationship between SPAD values and actual chlorophyll content was not strong in a semi-natural grassland, and SPAD measurements were poor predictors of total leaf chlorophyll content (all models R^2 = <0.5). Therefore, the use of a SPAD meter for plant communities, rather than for specific species, in heterogenous grasslands may not be applicable, although further research may be required.

The results were further explored to investigate whether specific sites had a greater influence on low R^2 values for trait prediction and why that might be, with the previous data collected over 2021. The data showed little consistency in the variation between sites or traits (Appendix D-2). There was some suggestive evidence that with increasing variation in a trait (i.e., greater range in trait values), the lower the predictive power of the model, however, the relationship was only significant for sward height (p = 0.0368), although all traits showed a negative trend. The predicted versus observed plots (Figure 5-6) also showed that few sites seemed to have extreme values in the structural traits and were far from the regression line, suggesting again that these extremes may reduce the models' predictive power. These identified sites tended to have more dominant tall growing grasses e.g., *Deschampsia cespitosa*, and potentially this influenced the model results. Siefert *et al.* (2015) demonstrated this, showing that intraspecific variation can influence 25% of variation within community level traits. Further exploration of the data also indicated that even within site variation may result in poor trait prediction, with certain sites having higher R^2 values one year, with low values the next. For example, the PLS-S2 model predicted AGB well for Eildon Hill in 2022 (R^2 = 0.925) but predicted AGB poorly in 2021 (R^2 = 0.490). When combining data from both years, the model predictive power was reduced further (R^2 = 0.168) (Figure 5-8).

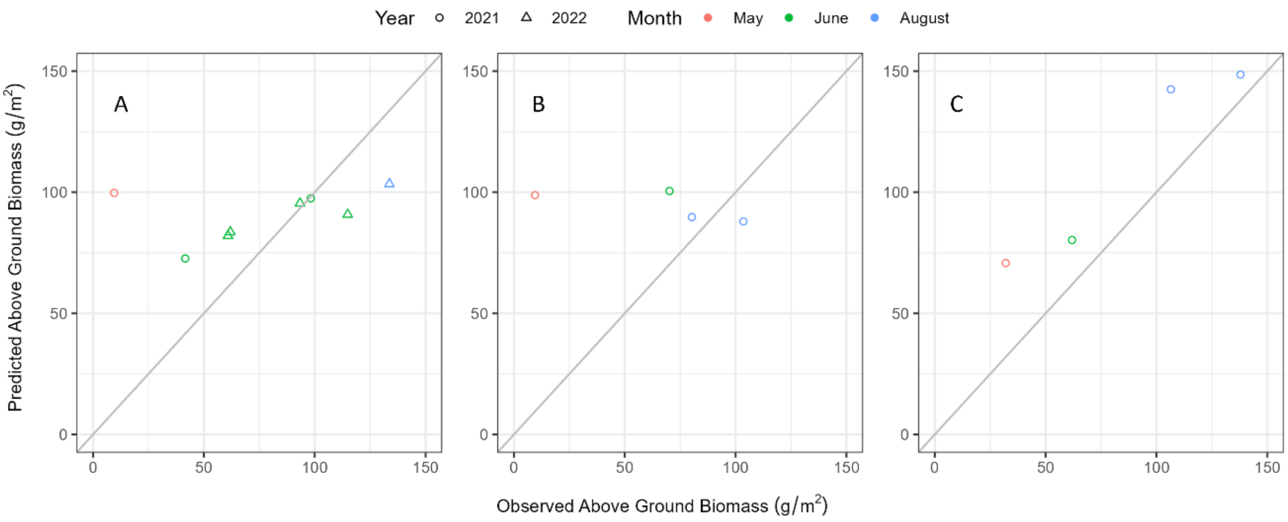


Figure 5-8. Predicted versus observed above ground biomass (g/m^2) at Eildon Hill from Sentinel-2 in 10 m resolution surface reflectance values and vegetation indices from a) 2021 and 2022 combined, b) 2021, and c) 2022.

Grasslands are varied by nature, showing quick responses to changes in conditions, climate, and management. This indicates that one-year trait prediction may be quite feasible, but another year a confounding factor has a greater influence on trait prediction which is reflected in the grassland community response, as explained by Thornley *et al*. (2023), and that collecting data across seasons invites even further variation into the data. Further research, plus personal observation, suggests that other topographical variables are important in influencing grassland characteristics e.g., elevation, slope, and aspect (Yin *et al*., 2019). It could be that predictive models for grassland community structural traits are best estimated with the inclusion of both spectral and topographical data. This knowledge could instead dictate what the data best be used for; potentially broad scale mapping may be limited from RS data but the difference in R^2 values per year at a specific site may be useful for monitoring purposes and assessing annual change.

The Micasense performed the best out of the three sensors, suggesting that trait prediction is spatially limited in species-rich grassland RS. This is supported in literature where Wang *et al*. (2018*a*) suggest the maximum pixel size for capturing the Spectral Variation Hypothesis is 10 cm, whilst Capolupo *et al*. (2015) note that the resolution should reflect the size of the desired trait to be estimated. For some traits PLS-S2 models only performed slightly worse than the PLS-MS models, particularly for AGB and fresh weight (Table 5-3). The differences seen here in the R^2 values between the sensor responses could be due to the variation seen on canopy versus leaf level. Due to the pixel size of Sentinel-2 and Planetscope, this data is extrapolated to canopy level, with the traits being measured as such e.g., AGB and fresh weight were recorded for the whole quadrat as a canopy level trait. However, due to the pixel size of the Micasense camera, this is more likely to be capturing leaf, and therefore species, level data. This could explain why the differences in the R^2 values are much larger for SMCP and sward height – which, although was averaged across canopy, the data was collected from individual plants – compared to AGB and fresh weight. The results from the models may also suggest that spatial resolution is a greater limiting factor where spectral resolution is not: Planetscope consistently performed the worst, despite having a higher spatial resolution than Sentinel-2 but has an inferior spectral resolution.

5.4.3 Limitations

Certain measurements were not considered in this study (e.g., species abundance) due, primarily, to time requirements. Unfortunately, due to the nature of some sites and access permissions, there was a difference in the number of sites that RS data could be collected at: Sentinel-2 and Planetscope imagery could be acquired for all sites, whereas UAV data could only be collected at seven sites. Scotland is also a country with highly changeable weather conditions; from May to August 2022 temperatures ranged from averages of 6.9 – 18.1 °C, whilst total rainfall varied from 80.6 – 121 mm (Met Office Climate Information Centre, 2023). Consequently, retrieval of satellite imagery (also limited by revisitation times) on the same day as *in situ* sampling was not always possible. Images that were cloud free over the site quadrats were acquired as close as possible to the field dates (Table 5-1). Similarly, poor weather conditions meant that the Micasense could not fly for some field dates.

Due to the reasons listed above, it is possible that these investigations in another subset of SRG sites might result in different findings. Further control variables may be needed to take into consideration, such as the specific timing of data collection (potentially limited to one survey at each site, rather than across a season), as well as matching management conditions as close as is possible. However, this reflects the nature of SRGs across Scotland and highlights the issues with up scaling these RS approaches nationally. Altering methods such as using a spectrophotometer for exact chlorophyll concentrations may be better represented in surface reflectance values, whereas measuring sward height and biomass with plate metres, for example, could be tested. This would allow comparisons to see if methodological changes help to improve results. It would be beneficial to explore these methods in more SRGs, as there was some evidence of success when focusing on specific sites. More complete field methodology recommendations are needed.

5.5 Conclusions

Species-rich grasslands are under mapped in Scotland but are listed as a priority habitat in conservation targets. RS data can be used to model and predict grassland community traits and species richness that vary between SRG classes. This information may have the potential to be utilised in improving mapping attempts on a wider scale.

This study explored RS applications in 11 species-rich grasslands in Scotland, across spatial and spectral scales. We investigated the Spectral Variation Hypothesis by modelling the relationship between species and spectral diversity. We also looked at trait retrieval of grassland community above ground biomass, grass fresh weight, sward height, and SPAD-measured chlorophyll-proxy across the various SRG types with predictive modelling. These common methods were integrated by utilising spectral diversity as a predictor variable in trait retrieval, an area that has had little exploration.

The results from this study lean toward greater questioning of the ability of prediction estimates from RS for supporting mapping of semi-natural SRGs. Although most of the R^2 values are low, there is still some variation explained by the RS data, especially for the structural trait estimation, that holds important information to be used (particularly when considering which predictor variables to input). It could simply be that there are too many considerations from these diverse and dynamic grassland habitats to accurately apply previous theories and common modelling approaches for SRGs. It appears that spectral variation in SRGs is not majorly linked to species counts but may more likely be representative of the species type, dominance, and environmental variance that is experienced regularly in these habitats.

Further data on potential confounding fixed and random factors needs to be considered in future prediction estimations in SRGs. This study does not wish to discount previous findings or theories surrounding the retrieval of grassland parameters from remotely sensed data, however, suggests that RS data alone may not yet be enough to predict certain plant community traits in SRGs with high accuracy, requiring inputs of additional data. We highlight that RS of SRGs is entirely community and site context dependent and potentially that will be a hinderance to progress in this area. When inclusive of multiple natural and semi-natural grassland classes, predicting diversity and community traits on a wide scale may fall short unless all possible dynamics are considered. It may be that initial *in situ* wide scale assessments per grassland community are needed before eventual groupings and consequent inferences for predicting and then mapping a broad range of grassland habitats by RS can occur.

5.5.1 Future Recommendations

It is still uncertain whether the retrieval and prediction of certain grassland parameters (both species diversity and community traits) can be used to aid the mapping of SRGs more widely. Time, labour, and financial costs will restrict the amount of data that can be collected. However, a set of ideal recommendations for RS of diverse grasslands can be derived.

The collection of further data should include species abundance, species functional group, number of flowering plants, and plant phenological stage to be considered as fixed effects (Fassnacht *et al.*, 2022; Schiefer *et al.,* 2021; Thornley *et al.* 2022; Wang *et al.,* 2018b). Data collection should make note of management practices, presence of disease, and weather conditions, specifically any unusual events as random effects (Thornley *et al.,* 2023; pers.obvs.). Temporal variation can be reduced by collecting data during one phenological stage, aligning any mapping attempts to this stage (Thornley *et al.,* 2022). The sensor and spatial resolution should be considered per trait estimation. Higher spatial resolution (cm scale) RS devices, such as UAVs should be used for species measured traits e.g., chlorophyll content, whereas lower spatial resolution (m scale) devices, such as certain satellites may be more applicable to measure canopy-level traits, such as above ground biomass (pers.obvs.). Where possible, influencing environmental determinants e.g., topographical variables should be included (Yin *et al.*, 2019).

These RS approaches were investigated to see whether retrieval and prediction of species diversity and community traits is possible across SRG classes. This was in attempt to assist future mapping of SRG classes which vary in their species and trait composition. What could be further explored is the use of clustering models here instead. Rather than use clustering models as an attempt to determine absolute number of species, for example, they appear more applicable to mapping biodiversity components especially regarding abundance and functional diversity estimations. This information may further inform SRG class prediction attempts. Future research can also look further into the integration of spectral diversity metrics in grassland community trait retrieval as some improvements in predictions were shown with this inclusion in this study.

References

Andreatta, D., Gianelle, D., Scotton, M. and Dalponte, M., 2022. Estimating grassland vegetation cover with remote sensing: A comparison between Landsat-8, Sentinel-2 and PlanetScope imagery. *Ecological Indicators*, *141*, p.109102.

Bardgett, R.D., Bullock, J.M., Lavorel, S., Manning, P., Schaffner, U., Ostle, N., Chomel, M., Durigan, G., L Fry, E., Johnson, D. and Lavallee, J.M., 2021. Combatting global grassland degradation. *Nature Reviews Earth & Environment*, *2*(10), pp.720-735.

Buisson, E., Fidelis, A., Overbeck, G.E., Schmidt, I.B., Durigan, G., Young, T.P., Alvarado, S.T., Arruda, A.J., Boisson, S., Bond, W. and Coutinho, A., 2021. A research agenda for the restoration of tropical and subtropical grasslands and savannas. *Restoration Ecology*, *29*, p.e13292.

Byrne, D., Moran, J., Maher, C. and Jones, G., 2018. Measure Handbook: Results-based Agri-environment Pilot Schemes – Species-rich Grassland and Wet Grasslands and Mosaic Habitat Suitable for Marsh Fritillary. Report prepared for the European Union, Agreement No. 07.027722/2014/697042/SUB/B2.

Capolupo, A., Kooistra, L., Berendonk, C., Boccia, L. and Suomalainen, J., 2015. Estimating plant traits of grasslands from UAV-acquired hyperspectral images: a comparison of statistical approaches. *ISPRS International Journal of Geo-Information*, *4*(4), pp.2792-2820.

Carbutt, C., Henwood, W.D. and Gilfedder, L.A., 2017. Global plight of native temperate grasslands: going, going, gone?. *Biodiversity and Conservation*, *26*(12), pp.2911-2932.

Conti, L., Malavasi, M., Galland, T., Komárek, J., Lagner, O., Carmona, C.P., De Bello, F., Rocchini, D. and Šímová, P., 2021. The relationship between species and spectral diversity in grassland communities is mediated by their vertical complexity. *Applied Vegetation Science*, *24*(3).

Dabrowska-Zielinska, K., Goliński, P., Jorgensen, M., Mølmann, J., Taff, G., Tomaszewska, M., Golińska, B., Budzynska, M. and Gatkowska, M., 2019. New methodologies for grasslands monitoring. *Proceedings of 23rd International Grassland Congress 2015-Keynote Lectures*.

De Keersmaecker, W., Van Rooijen, N., Lhermitte, S., Tits, L., Schaminée, J., Coppin, P., Honnay, O. and Somers, B., 2016. Species-rich semi-natural grasslands have a higher resistance but a lower resilience than intensively managed agricultural grasslands in response to climate anomalies. *Journal of Applied Ecology*, *53*(2), pp.430-439.

EEA, 2016. Mapping and assessing the condition of Europe's ecosystems: progress and challenges. *EEA Report No 3/2016*. Available at: https://www.eea.europa.eu/publications/mapping-europes-ecosystems

Fassnacht, F.E., Müllerová, J., Conti, L., Malavasi, M. and Schmidtlein, S., 2022. About the link between biodiversity and spectral variation. *Applied Vegetation Science*, *25*(1), p.e12643.

Féret, J.B. and Asner, G.P., 2014. Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. Ecological Applications, 24(6), pp.1289-1296

Friedl, M.A., McIver, D.K., Hodges, J.C., Zhang, X.Y., Muchoney, D., Strahler, A.H., Woodcock, C.E., Gopal, S., Schneider, A., Cooper, A. and Baccini, A., 2002. Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, *83*(1-2), pp.287-302.

Ge, Y., Atefi, A., Zhang, H., Miao, C., Ramamurthy, R.K., Sigmon, B., Yang, J. and Schnable, J.C., 2019. High-throughput analysis of leaf physiological and chemical traits with VIS–NIR–SWIR spectroscopy: a case study with a maize diversity panel. *Plant Methods*, *15*(1), pp.1-12.

Gholizadeh, H., Gamon, J.A., Helzer, C.J. and Cavender-Bares, J., 2020. Multi-temporal assessment of grassland α-and β-diversity using hyperspectral imaging. *Ecological Applications*, *30*(7), p.e02145.

Gholizadeh, H., Gamon, J.A., Zygielbaum, A.I., Wang, R., Schweiger, A.K. and Cavender-Bares, J., 2018. Remote sensing of biodiversity: Soil correction and data dimension reduction methods improve assessment of α-diversity (species richness) in prairie ecosystems. *Remote Sensing of Environment*, *206*, pp.240-253.

Giri, C., Zhu, Z. and Reed, B., 2005. A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. *Remote Sensing of Environment*, *94*(1), pp.123-132.

GLC., 2003. Global Land Cover 2000 database. European Commission, Joint Research Centre. Available at: http://www.gvm.jrc.it/glc2000.

Grüner, E., Astor, T. and Wachendorf, M., 2019. Biomass prediction of heterogeneous temperate grasslands using an SfM approach based on UAV imaging. *Agronomy*, *9*(2), p.54.

Hetzer, J., Huth, A. and Taubert, F., 2021. The importance of plant trait variability in grasslands: A modelling study. *Ecological Modelling*, *453*, p.109606.

Hijmans, R., 2023. *Raster: Geographic Data Analysis and Modeling*. R package version 3.6-21, https://rspatial.org/raster.

Homolová, L., Malenovský, Z., Clevers, J.G., García-Santos, G. and Schaepman, M.E., 2013. Review of optical-based remote sensing for plant trait mapping. *Ecological Complexity*, *15*, pp.1-16.

Imran, H.A., Gianelle, D., Rocchini, D., Dalponte, M., Martín, M.P., Sakowska, K., Wohlfahrt, G. and Vescovo, L., 2020. VIS-NIR, red-edge and NIR-shoulder based normalized vegetation indices response to co-varying leaf and Canopy structural traits in heterogeneous grasslands. *Remote Sensing*, *12*(14), p.2254.

Imran, H.A., Gianelle, D., Scotton, M., Rocchini, D., Dalponte, M., Macolino, S., Sakowska, K., Pornaro, C. And Vescovo, L., 2021. Potential and limitations of grasslands α-diversity prediction using fine-scale hyperspectral imagery. *Remote Sensing*, *13*(14), p.2649.

Jiménez-Jiménez, S.I., Marcial-Pablo, M.D.J., Ojeda-Bustamante, W., Sifuentes-Ibarra, E., Inzunza-Ibarra, M.A. and Sánchez-Cohen, I., 2022. VICAL: Global Calculator to Estimate Vegetation Indices for Agricultural Areas with Landsat and Sentinel-2 Data. *Agronomy*, *12*(7), p.1518.

Jin, Y., Yang, X., Qiu, J., Li, J., Gao, T., Wu, Q., Zhao, F., Ma, H., Yu, H. and Xu, B., 2014. Remote sensing-based biomass estimation and its spatio-temporal variations in temperate grassland, Northern China. *Remote Sensing*, *6*(2), pp.1496-1513.

JNCC., 2010. Handbook for Phase 1 habitat survey – a technique for environmental audit, JNCC, Peterborough, ISBN 0 86139 636 7.

Kilcoyne, A.M., Clement, M., Moore, C., Picton Phillipps, G.P., Keane, R., Woodget, A., Potter, S., Stefaniak, A. and Trippier, B., 2022. Living England: Technical User Guide. NERR108. Natural England.

Li, C., Wulf, H., Schmid, B., He, J.S. and Schaepman, M.E., 2018. Estimating plant traits of alpine grasslands on the Qinghai-Tibetan Plateau using remote sensing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(7), pp.2263-2275.

Li, C., Zhou, L. and Xu, W., 2021. Estimating aboveground biomass using Sentinel-2 MSI data and ensemble algorithms for grassland in the Shengjin Lake Wetland, China. *Remote Sensing*, *13*(8), p.1595.

Ludwig, A.D., Doktor, D., Goss, R., Sasso, S. and Feilhauer, H., 2022. The leaf is always greener on the other side of the lab: Optical in-situ indicators for leaf chlorophyll content need improvement for semi-natural grassland areas. *Ecological Indicators*, *143*, p.109424.

Lussem, U., Bolten, A., Menne, J., Gnyp, M.L., Schellberg, J. and Bareth, G., 2019. Estimating biomass in temperate grassland with high resolution canopy surface models from UAV-based RGB images and vegetation indices. *Journal of Applied Remote Sensing*, *13*(3), pp.034525-034525.

Magnificent Meadows., 2019. How to identify different types of grassland. Available from: http://www.magnificentmeadows.org.uk/assets/pdfs/How_to_identify_different_type s_of_grassland.pdf

Meng, B., Ge, J., Liang, T., Yang, S., Gao, J., Feng, Q., Cui, X., Huang, X. and Xie, H., 2017. Evaluation of remote sensing inversion error for the above-ground biomass of alpine meadow grassland based on multi-source satellite data. *Remote Sensing*, *9*(4), p.372.

Met Office Climate Information Centre., 2023. Areal values from HadUK-Grid 1km gridded climate data from land surface network. *[Online]*. Accessed 29th March 2023. Available at: https://www.metoffice.gov.uk/research/climate/maps-and-data/uk-and-regional-series

Milner, R. and Sharp, S., 2014. Molonglo River Reserve Procedures Manual: Monitoring Vegetation and Habitat Condition.

Möckel, T., Dalmayne, J., Prentice, H.C., Eklundh, L., Purschke, O., Schmidtlein, S. and Hall, K., 2014. Classification of grassland successional stages using airborne hyperspectral imagery. *Remote Sensing*, *6*(8), pp.7732-7761.

Möckel, T., Dalmayne, J., Schmid, B.C., Prentice, H.C. and Hall, K., 2016. Airborne hyperspectral data predict fine-scale plant species diversity in grazed dry grasslands. *Remote Sensing*, *8*(2), p.133.

Moeckel, T., Safari, H., Reddersen, B., Fricke, T. And Wachendorf, M., 2017. Fusion of ultrasonic and spectral sensor data for improving the estimation of biomass in grasslands with heterogeneous sward structure. *Remote Sensing*, *9*(1), p.98.

Muro, J., Linstädter, A., Magdon, P., Wöllauer, S., Männer, F.A., Schwarz, L.M., Ghazaryan, G., Schultz, J., Malenovský, Z. and Dubovyk, O., 2022. Predicting plant biomass and species richness in temperate grasslands across regions, time, and land management with remote sensing and deep learning. *Remote Sensing of Environment*, *282*, p.113262.

Muro, J., Schwarz, L., Männer, F., Linstädter, A. and Dubovyk, O., 2021, April. Modelling biomass and biodiversity in temperate grasslands with Sentinel-1, Sentinel-2 and Unmanned Aerial Vehicles. In *EGU General Assembly Conference Abstracts* (pp. EGU21-7232).

Palmer, M.W., Earls, P.G., Hoagland, B.W., White, P.S. and Wohlgemuth, T., 2002. Quantitative tools for perfecting species lists. *Environmetrics: The Official Journal of the International Environmetrics Society*, *13*(2), pp.121-137.

Pang, H., Zhang, A., Kang, X., He, N. and Dong, G., 2020. Estimation of the Grassland Aboveground Biomass of the Inner Mongolia Plateau Using the Simulated Spectra of Sentinel-2 Images. *Remote Sensing*, *12*(24), p.4155.

Peciña, M.V., Bergamo, T.F., Ward, R.D., Joyce, C.B. and Sepp, K., 2021. A novel UAV-based approach for biomass prediction and grassland structure assessment in coastal meadows. *Ecological Indicators*, *122*, p.107227.

Peng, Y., Fan, M., Bai, L., Sang, W., Feng, J., Zhao, Z. and Tao, Z., 2019. Identification of the best hyperspectral indices in estimating plant species richness in sandy grasslands. *Remote Sensing*, *11*(5), p.588.

Plantlife., 2018. Grasslands action plan. *[Online]*. Accessed 2nd January 2023. Available at: https://www.plantlife.org.uk/application/files/2315/3087/2058/Grasslands_action_plan_-_Plantlife.pdf

Polley, H.W., Yang, C., Wilsey, B.J. and Fay, P.A., 2019. Spectral heterogeneity predicts local-scale gamma and beta diversity of mesic grasslands. *Remote Sensing*, *11*(4), p.458.

Qin, Q., Xu, D., Hou, L., Shen, B. and Xin, X., 2021. Comparing vegetation indices from Sentinel-2 and Landsat 8 under different vegetation gradients based on a controlled grazing experiment. *Ecological Indicators*, *133*, p.108363.

Raab, C., Stroh, H.G., Tonn, B., Meißner, M., Rohwer, N., Balkenhol, N. and Isselstein, J., 2021. Mapping semi-natural grassland communities using multi-temporal RapidEye remote sensing data. In *Fine Resolution Remote Sensing of Species in Terrestrial and Coastal Ecosystems* (pp. 42-63). Routledge.

Reinermann, S., Asam, S. and Kuenzer, C., 2020. Remote sensing of grassland production and management—A review. *Remote Sensing*, *12*(12), p.1949.

Rocchini, D., Chiarucci, A. And Loiselle, S.A., 2004. Testing the spectral variation hypothesis by using satellite multispectral images. *Acta Oecologica*, *26*(2), pp.117-120.

Rocchini, D., Dadalt, L., Delucchi, L., Neteler, M. and Palmer, M.W., 2014. Disentangling the role of remotely sensed spectral heterogeneity as a proxy for North American plant species richness. *Community Ecology*, *15*(1), pp.37-43.

Rocchini, D., Marcantonio, M. And Ricotta, C., 2017. Measuring Rao's Q diversity index from remote sensing: An open source solution. *Ecological Indicators*, *72*, pp.234-238.

Rocchini, D., Salvatori, N., Beierkuhnlein, C., Chiarucci, A., De Boissieu, F., Förster, M., Garzon-Lopez, C.X., Gillespie, T.W., Hauffe, H.C., He, K.S. and Kleinschmit, B., 2021. From local spectral species to global spectral communities: A benchmark for ecosystem diversity estimate by remote sensing. *Ecological Informatics*, *61*, p.101195.

Rocchini, D., Santos, M.J., Ustin, S.L., Féret, J.B., Asner, G.P., Beierkuhnlein, C., Dalponte, M., Feilhauer, H., Foody, G.M., Geller, G.N. and Gillespie, T.W., 2022. The spectral species concept in living color. *Journal of geophysical research: Biogeosciences, 127*(9), p.e2022JG007026.

Rossi, C., Kneubühler, M., Schütz, M., Schaepman, M.E., Haller, R.M. and Risch, A.C., 2020. From local to regional: Functional diversity in differently managed alpine grasslands. *Remote Sensing of Environment*, *236*, p.111415.

Rossi, C., Kneubühler, M., Schütz, M., Schaepman, M.E., Haller, R.M. and Risch, A.C., 2021. Remote sensing of spectral diversity: A new methodological approach to account for spatio-temporal dissimilarities between plant communities. *Ecological Indicators*, *130*, p.108106.

Rossi, C., Kneubühler, M., Schütz, M., Schaepman, M.E., Haller, R.M. and Risch, A.C., 2022. Spatial resolution, spectral metrics and biomass are key aspects in estimating plant species richness from spectral diversity in species-rich grasslands. *Remote Sensing in Ecology and Conservation*, *8*(3), pp.297-314.

Rossi, M., Niedrist, G., Asam, S., Tonon, G., Tomelleri, E. and Zebisch, M., 2019. A comparison of the signal from diverse optical sensors for monitoring alpine grassland dynamics. *Remote Sensing*, *11*(3), p.296.

Rowland, C.S.; Marston, C.G.; Morton, R.D.; O'Neil, A.W. (2020). Land Cover Change 1990-2015 (25m raster, GB). NERC Environmental Information Data Centre. https://doi.org/10.5285/07b6e5e9-b766-48e5-a28c-5b3e35abecc0

Schiefer, F., Schmidtlein, S. and Kattenborn, T., 2021. The retrieval of plant functional traits from canopy spectra through RTM-inversions and statistical models are both critically affected by plant phenology. *Ecological Indicators*, *121*, p.107062.

Schmidtlein, S. and Fassnacht, F.E., 2017. The spectral variability hypothesis does not hold across landscapes. *Remote Sensing of Environment*, *192*, pp.114-125.

Scholtz, R. and Twidwell, D., 2022. The last continuous grasslands on Earth: Identification and conservation importance. *Conservation Science and Practice*, *4*(3), p.e626.

Schweiger, A.K. and Laliberté, E., 2022. Plant beta-diversity across biomes captured by imaging spectroscopy. *Nature Communications*, *13*(1), p.2767.

Schweiger, A.K., Schütz, M., Risch, A.C., Kneubühler, M., Haller, R. And Schaepman, M.E., 2017. How to predict plant functional types using imaging spectroscopy: linking vegetation community traits, plant functional types and spectral response. *Methods in Ecology and Evolution*, *8*(1), pp.86-95.

Shellswell, C.H. 2017. Is the rye-grass always greener? An evidence review of the nutritional, medicinal and production value of species-rich grassland. Plantlife. Available at: http://www.magnificentmeadows.org.uk/assets/pdfs/Is_the_rye-grass_always_greener_An_evidence_review.pdf

Siefert, A., Violle, C., Chalmandrier, L., Albert, C.H., Taudiere, A., Fajardo, A., Aarssen, L.W., Baraloto, C., Carlucci, M.B., Cianciaruso, M.V. and de L. Dantas, V., 2015. A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. *Ecology letters*, *18*(12), pp.1406-1419.

Soubry, I., Doan, T., Chu, T. and Guo, X., 2021. A systematic review on the integration of remote sensing and gis to forest and grassland ecosystem health attributes, indicators, and measures. *Remote Sensing*, *13*(16), p.3262.

Sullivan, C.A., Skeffington, M.S., Gormally, M.J. and Finn, J.A., 2010. The ecological status of grasslands on lowland farmlands in western Ireland and implications for grassland classification and nature value assessment. *Biological Conservation*, *143*(6), pp.1529-1539.

Süß, A., Danner, M., Obster, C., Locherer, M., Hank, T., Richter, K. and EnMAP Consortium, 2015. Measuring leaf chlorophyll content with the Konica Minolta SPAD-502Plus – Theory, Measurement, Problems, Interpretation. EnMAP Field Guides Technical Report, GFZ Data Services. DOI: http://doi.org/10.2312/enmap.2015.010

Swallow, K., 2016. Species-rich grassland: species identification and habitat classification. Accessed 25th October, 2022. Available at: https://naturebftb.co.uk/wp-content/uploads/2021/06/Species-rich-grassland-RSPB-presentation-23_11_20-PDF-CS.pdf

Tassi, A., Massetti, A. And Gil, A., 2022. The spectralrao-monitoring Python package: A RAO's Q diversity index-based application for land-cover/land-use change detection in multifunctional agricultural areas. *Computers and Electronics in Agriculture*, *196*, p.106861.

Thornley, R., Gerard, F.F., White, K. and Verhoef, A., 2022. Intra-annual taxonomic and phenological drivers of spectral variance in grasslands. *Remote Sensing of Environment*, *271*, p.112908.

Thornley, R.H., Gerard, F.F., White, K. and Verhoef, A., 2023. Prediction of Grassland Biodiversity Using Measures of Spectral Variance: A Meta-Analytical Review. *Remote Sensing*, *15*(3), p.668.

Török, P., Brudvig, L.A., Kollmann, J., N Price, J. and Tóthmérész, B., 2021. The present and future of grassland restoration. *Restoration Ecology*, *29*, p.e13378.

Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J.P., Veroustraete, F., Clevers, J.G. and Moreno, J., 2015. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties–A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *108*, pp.273-290.

Vickery, J.A., Tallowin, J.R., Feber, R.E., Asteraki, E.J., Atkinson, P.W., Fuller, R.J. and Brown, V.K., 2001. The management of lowland neutral grasslands in Britain: effects of agricultural practices on birds and their food resources. *Journal of Applied Ecology*, *38*(3), pp.647-664.

Wachendorf, M., Fricke, T. and Möckel, T., 2018. Remote sensing as a tool to assess botanical composition, structure, quantity and quality of temperate grasslands. *Grass and Forage Science*, *73*(1), pp.1-14.

Wang, R., Gamon, J.A., Cavender-Bares, J., Townsend, P.A. and Zygielbaum, A.I., 2018(a). The spatial sensitivity of the spectral diversity–biodiversity relationship: an experimental test in a prairie grassland. *Ecological Applications*, *28*(2), pp.541-556.

Wang, R., Gamon, J.A., Schweiger, A.K., Cavender-Bares, J., Townsend, P.A., Zygielbaum, A.I. and Kothari, S., 2018(b). Influence of species richness, evenness, and composition on optical diversity: A simulation study. *Remote Sensing of Environment*, *211*, pp.218-228.

Wang, Z., Ma, Y., Zhang, Y. and Shang, J., 2022. Review of Remote Sensing Applications in Grassland Monitoring. *Remote Sensing*, *14*(12), p.2903.

Wilsey, B., 2021. Restoration in the face of changing climate: importance of persistence, priority effects, and species diversity. *Restoration Ecology*, *29*, p.e13132.

Yin, J., Feng, Q., Liang, T., Meng, B., Yang, S., Gao, J., Ge, J., Hou, M., Liu, J., Wang, W. and Yu, H., 2019. Estimation of grassland height based on the random forest algorithm and remote sensing in the Tibetan Plateau. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, pp.178-186.

Zhang, B., Zhang, L., Xie, D., Yin, X., Liu, C. and Liu, G., 2015. Application of synthetic NDVI time series blended from Landsat and MODIS data for grassland biomass estimation. *Remote Sensing*, *8*(1), p.10.

Zhang, H., Sun, Y., Chang, L., Qin, Y., Chen, J., Qin, Y., Du, J., Yi, S. and Wang, Y., 2018. Estimation of grassland canopy height and aboveground biomass at the quadrat scale using unmanned aerial vehicle. *Remote Sensing*, *10*(6), p.851.

Zhang, L., Gao, H. and Zhang, X., 2023. Combining Radiative Transfer Model and Regression Algorithms for Estimating Aboveground Biomass of Grassland in West Ujimqin, China. *Remote Sensing*, *15*(11), p.2918.

Zhang, Y.W., Wang, T., Guo, Y., Skidmore, A., Zhang, Z., Tang, R., Song, S. and Tang, Z., 2022. Estimating community-level plant functional traits in a species-rich alpine meadow using UAV image spectroscopy. *Remote Sensing*, *14*(14), p.3399.

Zhao, Y., Sun, Y., Chen, W., Zhao, Y., Liu, X. and Bai, Y., 2021(a). The potential of mapping grassland plant diversity with the links among spectral diversity, functional trait diversity, and species diversity. *Remote Sensing*, *13*(15), p.3034.

Zhao, Y., Sun, Y., Lu, X., Zhao, X., Yang, L., Sun, Z. and Bai, Y., 2021(b). Hyperspectral retrieval of leaf physiological traits and their links to ecosystem productivity in grassland monocultures. *Ecological Indicators*, *122*, p.107267.

Zou, X., Zhu, S. and Mõttus, M., 2022. Estimation of canopy structure of field crops using sentinel-2 bands with vegetation indices and machine learning algorithms. *Remote Sensing*, *14*(12), p.2849.

Chapter 6. A Citizen Science Tool for Mapping Species-Rich Grasslands: Uniting the Public with Remote Sensing for Biodiversity Conservation

Status: In preparation for submission to Nature Communications.

Abstract

As a society we continuously miss global targets that will help us address the ecological crisis. Collaboration is required to provide an integrative approach for improving biodiversity monitoring to address this crisis. Proposals have been made to combine two of the largest applications in nature and landcover monitoring, citizen science (CS) and remote sensing (RS), both of which can be used to improve spatial reach of monitoring initiatives. CS can reduce financial constraints and increase knowledge exchange, whilst RS can help access isolated regions and improve the temporal scale of monitoring. However, few studies explore the combination of CS and RS due to data quality concerns and discipline-specific knowledge. This chapter aimed to address this gap by combining CS and RS in a novel biodiversity monitoring tool. A citizen science survey, Ecosystem Explorers, was co-created with Butterfly Conservation to help locate species-rich grasslands (SRGs) in Scotland; a habitat that has seen large declines over the last 60 years. This interdisciplinary project was planned to engage stakeholders in biodiversity monitoring attempts utilising open science practices and facilitating the combination of CS and RS. Participants were recruited to confirm areas of SRGs that had been predicted by a habitat classification model, applied to Sentinel-2 satellite imagery. Data was also provided by NatureScot, Plantlife, and the Botanical Society of Britain and Ireland on recently confirmed SRG locations. Model predicted SRG locations were cross-referenced with citizen ground-truthed SRG locations and the level of agreement was analysed. RS data was successfully utilised in a CS project initiating interdisciplinary methods, whilst collaboration and stakeholder engagement resulted in the creation of an interdisciplinary tool. Through implementing open science practices, the finished project is predicted to have a high open score of 0.92. There was a low overall agreement between the model predicted and citizen ground-truthed SRG locations (17.65%). The difference in the number of model predicted and citizen ground-truthed SRG locations was significantly different to the theoretical expected equal split of model predicted and citizen ground-truthed SRG locations ($p = 0.0003$). The level of agreement between the model predicted and citizen ground-truthed SRG locations was significantly different across SRG classes: acid, calcareous, marshy, and neutral ($p = 0.017$). However, the level of agreement between the model predicted and citizen ground-truthed SRG locations was not significantly different across data providers, participant experience, or participant confidence level ($p = 0.114$, $p = 0.706$, $p = 0.900$ respectively). The results suggest there are difficulties in using the current combined tool for wide-scale mapping of species-rich grasslands in Scotland due to the high disagreement between the model predictions and citizen scientists. However, the methodology derives an open science approach that could be further enhanced to improve global biodiversity monitoring, that is context specific. The results support the involvement of stakeholder and community engagement in biodiversity conservation research.

6.1 Introduction

6.1.1 Global Biodiversity Monitoring

The biodiversity crisis, defined by Rahbek (2012) as the “rapid loss of species and … degradation of ecosystems”, is recognised as a global issue and threat to sustainable living (De Prins, 2022; Purnomo *et al.,* 2020; Singh, 2002; Steffen, 2003; Western, 1992). This awareness has led to the creation of the UN Millenium Development Goals, followed by the SDGs and Aichi Targets (Brooks *et al.,* 2015; Sachs, 2012) as well as local, national, and international initiatives to mitigate and/or reverse the damage caused by human activities. A major requirement in meeting targets designed to halt or reverse biodiversity loss is comprehensive environmental monitoring. This monitoring enables researchers to identify changes in species’ populations and environmental functioning, the potential causes of these changes, and the solutions to halt and reverse them. However, environmental monitoring is limited by a variety of challenges, such as funding limitations, researcher availability, and site accessibility, which reduce the capacity for complete monitoring (Anderson, 2018). This has led to large spatial, temporal, and taxa-specific gaps in coverage (Xu *et al.,* 2021). These excluded species, habitats, regions, and minority groups in society often experience the effects of the ecological crisis most greatly. For example, the tropics, where monitoring programmes are not nearly as abundant as in the Global North, have the highest rates of biodiversity loss and the communities found here are much more dependent on biodiversity resources for their livelihood (Roe *et al.,* 2019).

Globally, grassland habitats are often overlooked in terms of monitoring, management, and conservation; this may prove problematic, due to their carbon sequestration capacity and role as a potential climate change mitigator (Bardgett *et al.,* 2021; Carbutt *et al.,* 2017), as well as their importance as key habitats for a range of species. Grassland degradation has directly contributed to the current decline in insect populations, as well as climate change, impacting the ecosystem services they provide (including pollination, nutrient cycling, and pest control) (Wilson and Fox, 2021). Species-rich grasslands (SRGs) in the UK make up less than 1% of the UK landscape and their loss is coupled with declines in dependent insect species, such as that of the Northern Brown Argus butterfly, *Aricia artaxerxes*, whose population has more than halved in the last 50 years (Natural England, 2020). Without consistent monitoring and management these fragile ecosystems and associated species will continue towards their extinction.

Issues facing environmental monitoring include the need for an increase in data collation capacity, multiple stakeholder engagement, and cross-discipline approaches (Kühl *et al.*, 2020). As indicated in Chapter 2, a potential approach to meeting this challenge is the application of the principles of the OS movement, which facilitates increased collaboration in environmental monitoring and creating a more accessible scientific community and research process (Hecker *et al.*, 2018). An OS tool that has increased in its use over the last two centuries is the use of Citizen Science – the generally accepted definition of which is the participation of untrained volunteers in scientific research, largely for data collection (Haklay, 2021).

6.1.2 Citizen Science in Environmental Monitoring

Citizen science (CS) has some of its largest applications in environmental monitoring, where volunteers can increase the amount of data collected as well as the number of habitats and species covered (de Sherbinin *et al.,* 2021). CS projects provide environmental scientists with a low-cost method that can increase the spatial and temporal scale of data collection, whilst also engaging the public and stakeholders with the potential for policy implementation (Fritz *et al.*, 2019). There are previous concerns in CS projects regarding participants' knowledge of biodiversity and monitoring. However, these concerns are being addressed with the implementation of simple methods, volunteer training, and continued feedback in projects. This has been shown to improve accuracy amongst data collectors with projects finding high agreement between researchers and volunteers (Kosmala *et al.,* 2016).

In fact, CS has the potential to contribute to approximately 33% of the current SDGs, with its largest and potential contributions in 1) habitat restoration initiatives (SDG 15), 2) making cities sustainable (SDG 11), 3) promoting human wellbeing (SDG 3), and 4) water monitoring and sanitation initiatives (SDG 6) (Fraisl *et al.*, 2020). Fritz *et al.* (2019) outlined the ways in which citizen science can contribute to the SDGs through a combination of greater temporal and spatial reach, multi-dimensionality across SDGs and providing richer/detailed data, and data management for accessibility (Figure 6-1).

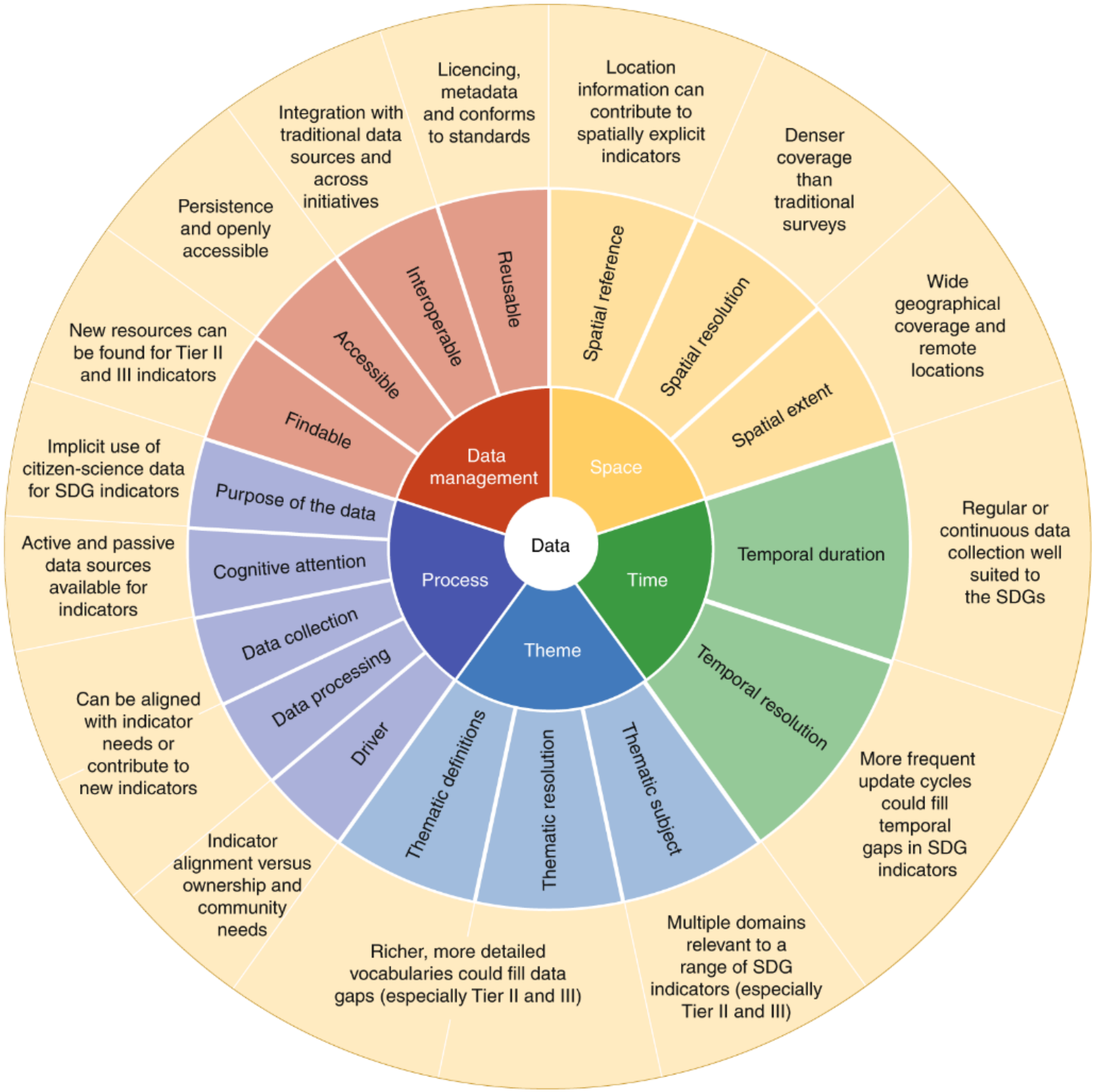


Figure 6-1. The potential contribution of citizen science to the Sustainable Development Goals. Source: Fritz *et al*. (2019).

There is an opportunity for CS projects to increase the impact on biodiversity conservation outside of increased data collection, particularly in relation to participant understanding. Participant awareness and knowledge of the project's target species or habitat may be improved by taking part in a project. For example, Greving *et al.* (2022) found that, in an urban bat ecology CS project, participation resulted in a greater understanding of urban bat ecology. As such, CS participation could influence conservation efforts (for insects in particular, which are often ignored due to dislike and unfamiliarity). This could be because increased awareness and ownership may result in both engagement with public campaigning and changes to potentially harmful behaviours (Saunders *et al.*, 2020). Furthermore, CS may change environmental policy and lead to increased protection, or just greater acceptance, of overlooked species (Adler *et al.*, 2020).

6.1.3 Remote Sensing in Citizen Science

Although recognition of the role of CS as a robust scientific method continues to increase, areas for improvement remain, including refining data integration and modelling approaches, reducing observer bias, and widening scientific communication of research outputs and results (Johnston *et al.*, 2022). One area that is frequently highlighted for increased integration with CS is the use of remote sensing (RS) technologies (defined as observing the Earth with the use of satellites or drones) in surveys (Lee *et al.*, 2020). For example, CS projects have the potential to address multiple Essential Biodiversity Variables (EBVs), such as monitoring changes in species population trends, genetic diversity, and ecosystem functioning, however, most typically focus on just one target (e.g., species-based projects) (Turak *et al.*, 2017; Wetzel *et al.*, 2018). The use of RS in CS could address this concern by being applicable to wider scale ecosystem-based monitoring (for example, observing grasslands and associated species), therefore, targeting multiple EBVs at once.

Open source RS imagery, although increasingly accessible, is a resource that has not been fully utilised in CS. This is largely due to misconceptions in Earth observation about the validity of CS data, as well as concerns regarding the transferability of Earth observation to biodiversity monitoring (Schulte to Bühne and Pettorelli, 2018). However, the potential for combining RS data in a CS project is increasingly being explored, with projects starting to demonstrate success. These include the syncing of tree phenological timings *in situ* with near-surface RS imagery (Kosmala *et al.*, 2016), and the identification of archaeological artefacts within RS LiDAR imagery (Lambers *et al.*, 2019).

There are multiple ways in which citizen scientists and RS can interact, including: ground-truthing models, confirming RS outputs, providing local knowledge, or filling information gaps (Boyd *et al.*, 2022). An example of this is seen in the project by Domingo-Marimon *et al.* (2020) where phenological maps produced from satellite data were provided to members of the public to consolidate the outputs. The scale of environmental monitoring can be extended even further through the combination of methods. This combination will allow inaccessible regions to be reached and cover larger areas at a quicker rate than even the use of many citizens can achieve, further maximising the potential of each method (Stephenson, 2020). However, as Stephenson (2019) addresses, the RS technology used must be appropriate for the monitoring design in question rather than the other way around.

The impact of combining CS and RS can be further explored with the potential for project interaction, which can be achieved more easily with the practice of OS (Lee *et al.*, 2020). For example, if a CS project, that is not defined by the use of RS, is collecting data that includes habitat identification (as an example), this could feed into RS projects that are looking at land cover changes, in spite of this not being a primary aim of the CS project. This can largely benefit the provision of reference datasets for RS products (See *et al.*, 2022). It does appear that the current combination of CS and RS is heavily focused on the validation and calibration of RS data (See *et al.*, 2016). However, these projects may be further improved by exploring the increase in co-design of CS and RS projects, with local communities and stakeholders helping to inform project direction too.

The outcomes of well-designed combined CS and RS need to be two-fold. In addition to the use of citizen scientists in RS projects increasing the spatial reach and data collection capacity, projects must also ensure advantages of participation for the citizen scientists. With RS having large applications in Earth observation and, increasingly, in biodiversity monitoring, interaction with RS products provides the opportunity for the public to engage with the pressures the Earth is experiencing and feel like an active contributor to the resolution of these. There are psychological barriers that are experienced by the public regarding the ecological crisis and climate change; research shows that these are difficult concepts for people to grasp, as it is not an immediate, tangible, and personal threat that they face in their day to day lives (Gifford, 2011; Scannell and Gifford, 2013). There is also evidence that behavioural change comes from self-persuasion i.e., for any action to occur an individual must make their own choice (De Meyer *et al.,* 2020). This may be more likely to occur through their direct experiences and from learning by others' examples (which CS would provide the opportunity for), rather than being told what they should or should not do. Involvement in CS and RS may be a way to improve Eco-pedagogy – the establishment of ecological consciousness from direct involvement or "confrontation" leading to action (Dunkley, 2018). If this engagement is followed by pro-environmental behavioural change, it has the potential to benefit both the planet, through action, and people, through improved wellbeing and resulting future sustainability (Zawadzki *et al.,* 2020).

Combining RS and CS methodologies has great potential, with advantages ranging from increased spatial reach and validation of RS projects, to heightened participant scientific democratisation and biodiversity awareness. However, this combination is still underexplored and is not without challenges. These challenges have broadly been identified as 1) quality assurance and trust of CS data, 2) complex RS data and lack of easy-access technology, and 3) the value provided by the participants and how to support them (Mazumdar *et al.,* 2017). Hopefully, these challenges will be resolved with the rising union of CS and RS, especially where project efforts outline how they will address them.

6.1.4 Maximising "Openness" in Citizen Science: Collaboration in Design

Collaboration and co-creation are becoming increasingly important in the development of CS projects (Tan *et al.,* 2022). This interactive process ensures that CS projects are designed to target the needs of society and gaps in scientific research that are identified by the parties who will be most affected by the project. This collaboration can also address concerns of knowledge gaps in interdisciplinary approaches, such as those seen when combining Earth science with conservation science, for example (Gunnell *et al.,* 2021). It also enables projects to become widely cross-disciplinary in their applications and achieve greater impact amongst science and society through this approach (Senabre Hidalgo *et al.,* 2021). This level of public and stakeholder engagement also ensures feedback to assess the success of the project. To ensure project targets are suitably identified and reached, communication via multiple stakeholders is a necessity in project creation.

While CS projects are set up with multiple goals at their respective cores, these are often not reached. For example, Theobald *et al.* (2015) found that, although 376 projects had an aim to contribute to scientific knowledge, just 47 of the biodiversity CS projects were published in a peer-reviewed journal. To ensure the goals of these CS projects are realised, intentional fit-for-purpose design is required by creating a workflow that targets each aim (Parrish *et al.,* 2018). Golumbic *et al.* (2020) cross-analysed five highly successful CS projects and found attaining their goals was associated with i) a user-friendly platform, ii) simple tasks with relevant training, iii) recruiting existing interested participants and those in need, iv) active communication with other participants and researchers, and v) transparency and availability of results. The likelihood of success when creating a CS project can be associated with these findings as well as through following the 10 principles of CS, which highlights the need for communication, transparency, mutually beneficial goals to participants and researchers, and high quality data for scientific advancement (Table 6-1) (Robinson *et al.,* 2018; Vohland *et al.,* 2021).

Table 6-1. The 10 principles of Citizen Science. Source: ECSA, 2015.

Citizen Science Principles	
1. Citizen science projects actively involve citizens in scientific endeavour that generates new knowledge or understanding. Citizens may act as contributors, collaborators or as project leaders and have a meaningful role in the project.	6. Citizen science is considered a research approach like any other, with limitations and biases that should be considered and controlled for. However, unlike traditional research approaches, citizen science provides opportunity for greater public engagement and democratisation of science.
2. Citizen science projects have a genuine science outcome. For example, answering a research question or informing conservation action, management decisions or environmental policy.	7. Citizen science project data and metadata are made publicly available, and, where possible, results are published in an open-access format. Data sharing may occur during or after the project unless there are security or privacy concerns that prevent this.
3. Both the professional scientists and the citizen scientists benefit from taking part. Benefits may include the publication of research outputs, learning opportunities, personal enjoyment, social benefits, satisfaction through contributing to scientific evidence, for example, to address local, national, and international issues, and through that, the potential to influence policy.	8. Citizen scientists are acknowledged in project results and publications.
4. Citizen scientists may, if they wish, participate in multiple stages of the scientific process. This may include developing the research question, designing the method, gathering and analysing data, and communicating the results.	9. Citizen science programmes are evaluated for their scientific output, data quality, participant experience and wider societal or policy impact
5. Citizen scientists receive feedback from the project. For example, how their data are being used and what the research, policy or societal outcomes are.	10. The leaders of citizen science projects take into consideration legal and ethical issues surrounding copyright, intellectual property, data-sharing agreements, confidentiality, attribution, and the environmental impact of any activities.

6.2 Aims and Objectives

The research in this chapter ultimately brings together the various approaches that have been undertaken in the previous chapters to address the initial thesis aims. Ultimately, two main aims are being addressed here, the thesis RQ5: Can an interdisciplinary, open science citizen science project be created, utilising remote sensing outputs for habitat monitoring?; and the case study specific question: Can citizen science data validate the outputs of remote sensing models to identify species-rich grasslands for vulnerable species protection?

The specific research questions asked were:

i.) Where can open science practices be implemented in the creation of the combined tool?

ii.) How can remote sensing data be utilised in a citizen science survey to create an interdisciplinary tool?

iii.) Can the tool accurately locate species-rich grasslands and target species habitat?

iv.) Does participant knowledge and experience affect the success of predicting species-rich grasslands?

6.3 Methods

The methods throughout this chapter follow the creation and initiation of the CS survey, Ecosystem Explorers, with Butterfly Conservation, that invited members of the public to help map SRGs across Scotland (section 6.3.1). The RS model outputs from chapter 4, of predicted locations of SRGs, were then utilised in the resulting CS survey, allowing discipline integration to occur and the combined tool to be created (section 6.3.2). This allowed the exploration of how the thesis research questions were met, with the methods for this evaluation outlined. Following on from this, the specific accuracy of the tool for mapping SRGs and locating *A. artaxerxes* habitat was then tested utilising the data collected by citizen scientists and other stakeholders and participants, addressing the case-study specific research aims. Further introduction of other sources of data (from NGOs and participants outside of the CS survey), discussed below, was necessary due to lower-than-expected participation rates in the Ecosystem Explorers survey, as well as for providing further information on participant skills (section 6.3.2 and 6.3.3). From this, analysis could investigate the specific requirements for success of combining RS and CS and for mapping SRGs on a wide scale (6.3.4).

6.3.1 Developing the Citizen Science Project

6.3.1.1 Co-Creation

To ensure that the data generated would be of use to multiple stakeholders, and that local expert knowledge ascertained which priority conservation areas to pursue, the CS survey was co-designed in partnership with NGO Butterfly Conservation. This iterative approach aimed to ensure that the data generated by the project addressed biodiversity gaps and targets beyond those specific to this thesis, in addition to providing access to a pool of available volunteers willing to participate in the surveys.

Meetings were held with members of Butterfly Conservation at three points throughout the project: an initial meeting in 2020 was set up to highlight specific vulnerable species in the project such as *A. artaxerxes*. Various survey protocols were co-designed with adaptations on Butterfly Conservation's current transect, butterfly, and egg surveying methodologies to ensure consistency. A follow-up meeting was held before the initial survey season in 2021 to discuss extra data that could be collected regarding site condition and the presence of priority species. A third meeting after the first survey season was held to outline the aims of the CS survey and ensure sufficient data capture that would meet the needs of this thesis and Butterfly Conservation's targets. Informal communication has been used to provide updates, refine objectives, allow promotion of the project, and disseminate results. The resulting CS survey was called 'Ecosystem Explorers'.

6.3.1.2 Identifying Suitable Species-Rich Grassland Classifications and Indicators

The SRGs that were investigated are described as acid, neutral, calcareous, and seasonally wet/marshy grasslands, as used when describing habitats in Butterfly Conservation's transect surveys (see chapter 3 for greater descriptions). A final class (coastal grasslands) was also included, not defined in classification schemas but found to be relevant due to the different species presence found *in situ*. These classes are defined by their abiotic factors (such as soil pH, moisture, and nutrient content) arising to different species occurrences, ultimately determining the habitat class.

Most habitat classifications are performed using visual assessment, for example, Phase 1 and NVC surveys, and the CS survey incorporated this methodology. Key indicator species were chosen per habitat class based on field guides and literature (from the Field Studies Council Guides, the Magnificent Meadows project, and the Species Recovery Trust, see Appendix E-1) that outline these class descriptions for determining the grassland types of the survey. For example, Wild Thyme that is associated with calcareous grasslands or Common Tormentil associated with acidic grasslands (Table 6-2). These key indicator species were used as guides to help participants determine habitat class and found in an accessible and downloadable PDF document (see Appendix E-2). It is important to note that many indicator species will be present in multiple grassland classes due to the transitional and graduating nature of grassland habitats. The specific indicator species were chosen for each class as are most often found in their associated grassland type and in conjunction with the other indicators in that class. The aim was to determine the class as best as possible based on the composition of key indicator species found, and at what frequencies, using own judgement and justification.

Table 6-2. Key indicator species used to visually determine grassland habitat classes in a citizen science survey. This list is not exhaustive, but species were used as a baseline for determining species-rich grasslands along with a list of resources available. Indicator species adapted from the Field Studies Council guides, the Magnificent Meadows project, and the Species Recovery Trust.

Grassland Class	Key Indicator Species		
	Graminoids	Herbs	Lepidoptera
Acid	*Deschampsia flexuosa, Nardus stricta, Festuca ovina, Muhlenbergia rigens, Agrostis capillaris*	*Galium saxatile, Potentilla erecta, Lathyrus linifolius, Rumex acetosella,*	*Coenonympha pamphilus, Maniola jurtina, Lycaena phlaeas*
Neutral	*Dactylis glomerata, Holcus lanatus, Cynosurus cristatus, Anthoxanthum odoratum, Arrhenatherum elatius*	*Centaurea nigra, Lathyrus pratensis, Ranunculus acris, Cirsium arvense, Rhinanthus minor, Plantago lanceolata*	*Maniola jurtina, Polyommatus icarus, Aphantopus hyperantus, Odezia atrata, Coenonympha pamphilus*
Calcareous	*Brizia media, Sesleria caerulea, Helictotrichon pratense*	*Helianthemum nummularium, Thymus polytrichus, Galium verum, Sanguisorba minor, Linum catharticum*	*Aricia Artaxerxes, Zygaena filipendulae, Cupido minimus, Melanargia galathea, Pyrgus malvae*
Marshy	*Molinia caerulea, Juncus sp., Carex sp.*	*Caltha palustris, Cirsium palustre, Filipendula ulmaria, Angelica sylvestris, Valeriana dioica, Achillea ptarmica, Carum verticillatum*	*Euphydryas aurinia*
Coastal	*Festuca ovina/rubra, Ammophila sp., Poa pratensis*	*Armeria maritima, Silene uniflora, Lotus corniculatus, Thymus polytrichus*	*Lycaena phlaeas, Cupido minimus*

6.3.1.3 Identifying Key Food Plants and Butterfly Species

In line with Butterfly Conservation's goals, the survey included butterfly transect and egg surveys, particularly for the key butterfly species (Northern Brown Argus, *A. artaxerxes*), and report the presence of any key food plants (in the case of *A. Artaxerxes,* the caterpillar food plant, *H. nummularium*).

6.3.2 The Survey: Ground-Truthing Predicted Species-Rich Grasslands using Citizen Science Methods

6.3.2.1 Using Classification Maps for Directed Surveying

A species-rich grassland classification model was created by integrating previously collected remote sensing and environmental data from 2021 (outlined in chapter 3). Spectral data was collected from satellites, including Sentinel-2 (S2) and PlanetScope, as well as data from Unoccupied Aerial Vehicles (UAVs). Environmental variables were collected as predictors of habitat classification, which included AGB, sward height, soil pH, soil moisture, soil bulk density, soil texture, and species richness. The classification model was created in R (methodology outlined in chapter 3) and applied to the S2 satellite imagery to further predict areas of different grassland classes (neutral, acid, calcareous, marshy, and coastal grasslands) across Scotland.

Classification maps were produced from the model, predicting areas of SRGs. Initial maps were provided to participants as 5 km^2 regions, georeferenced by OS 5 km^2 grids and location names of features in the area, such as water bodies, towns, and hills. Aerial images (at 25 cm spatial resolution) of the corresponding area were also provided to further help participants. An example of this is seen in Figure 6-2. Random points were generated (due to the large spatial scale covered by the model outputs) across the classification maps in ArcGIS Pro, to provide a greater number of survey locations of potential SRGs intended for the public. Both classification maps and the coordinates of potential locations of identified grassland habitats were used as the outputs to inform the CS survey. Participants could choose a survey location by either using the coordinates from the list provided or decide on their own location by using the maps instead.

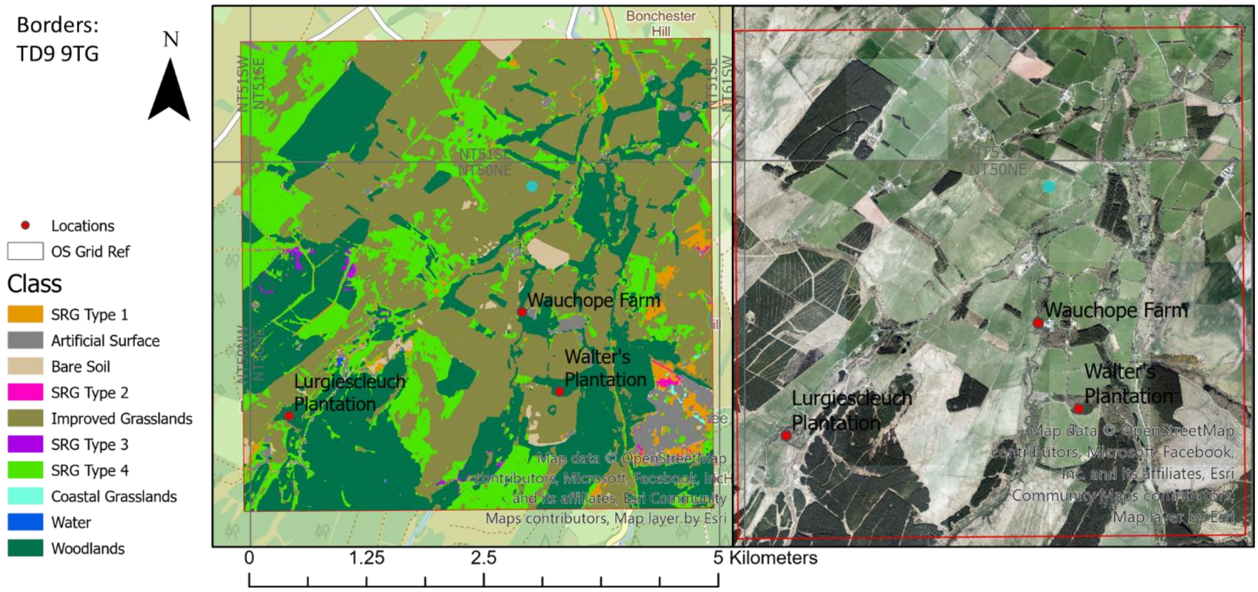


Figure 6-2. Habitat prediction map output from remote sensing data, supplied to participants of the Ecosystem Explorers citizen science survey, for ground-truthing predictions of species-rich grasslands.

The model outputs were used to guide the participants to the identified locations where they were able to confirm or deny the presence of the model predicted grassland class. Data was supplied by participants who had been directed by the classification maps or location points to a predicted area of SRG. Participants conducted the surveys outlined by the methodology below and the initial GPS location provided (pixel) was used for comparison with the classification map.

6.3.2.2 The Habitat and Species Presence Assessments

Whilst at the sites, citizen scientists were given a data collection sheet (see Appendix E-3) to fill out, including information on site conditions, management plans, prior survey experience and species identification knowledge, habitat classification, and presence of key plant and butterfly species. To ensure consistency from the various data collection sources and methods, the data was standardised by requiring specific information from participants and providers, summarised as data that was critical for analysis and data that provided extra value-added information but did not necessarily help confirm the model outputs and abilities of participants (Table 6-3).

Table 6-3. Type of data needed for analysis of habitat prediction model success and effect of participant ability versus data that provided extra value-added information.

Critical for analysis	Extra value-added
Site location point	Species presence and abundance
Observed species-rich grassland class	Site and weather conditions
Participant experience level	Observed lepidopteran species (excluding *Aricia artaxerxes*)
Participant confidence level	Species richness and sward height
Rock-rose presence	Participant age range
Aricia artaxerxes egg or adult sightings	

On site, participants conducted transect surveys. Transects were randomly located on the site but could vary in length due to site constraints. However, transects were no longer than 1 km. Participants observed the habitat whilst walking the transect and recording any butterfly species within 5 m either side of the transect. This was to align with Butterfly Conservation's Butterfly Transect Guides. Along the transect, three quadrat surveys (the length of an A4 piece of paper squared) were carried out at the beginning, middle, and end of the transect. This allowed participants to not require any field equipment, increasing the accessibility of the project. The participants recorded species richness, sward height, and where possible listed the species present, with the help of ID guides and apps (see Appendix E-1 and Appendix E-2), in the three quadrats. Participants were also asked to take any pictures of the sites. No destructive measurements were to be taken due to permissions needed, lack of experience, and time requirements. The observations on site and species presence were used for the participants to determine the habitat or grassland class. Whilst conducting the transects, participants were asked to note the presence of any *H. nummularium*. If *H. nummularium* was found, a *A. artaxerxes* egg survey was to be conducted by recording the presence and GPS location of any eggs located. Presence and count of any adult *A. artaxerxes* was also to be noted whilst participants were out surveying (see Appendix E-3).

Data was collected through multiple means to give participants options suited to them. For example, printable survey forms were available on the web platform that participants could take into the field and then transcribe later, on the web platform. The other option was for participants to access the project in field through the citsci app and enter data virtually whilst out surveying, to reduce issues with printing and save time transcribing. Other participants provided data via email or handed data collection sheets to me at Ecosystem Explorers survey days. The data on the web platform was exported to an excel sheet and collated with externally provided data (see section 5.3.3) for analysis.

6.3.2.3 Participant Information

Little demographic data of participants was collected due to ethical restraints. However, participants were asked to record their age range, their previous plant identification and biodiversity surveying experience, and their confidence in their habitat classification assessment (Appendix E-3). Participants noted their previous identification and surveying experience as either 'None/Moderate/Advanced'. They also rated their confidence in their habitat classification assessment using a Likert scale between 1 – 5, ranging from ‘no confidence’ to ‘very confident’ respectively. This was both for quality assurance of the data and providing insights into the people who choose to participate in biodiversity monitoring surveys. This way, it was possible to assess what age groups are more likely to participate, and if certain age groups have greater knowledge, experience, and confidence in biodiversity assessments. This is important in a time where increased use of tech devices has a negative impact on nature connectedness at crucial life stages (e.g., teenage years), whilst access to botanical courses has decreased in formal education (Price *et al.,* 2022; Stroud *et al.,* 2022). These factors have led to a decline in plant literacy and will have severe consequences for biodiversity conservation as well as people’s enjoyment of nature (Barrable *et al.,* 2021; Thomas *et al.,* 2022).

6.3.2.4 The Citizen Science Platform and Data Collation

An online web platform (https://www.citsci.org/projects/ecosystem-explorers) was designed for the survey using the project builder function on citsci.org. This is a commonly used site to develop and promote CS projects that operates under an OS approach, where project development and participation is free, and data is available under an open creative commons license (Lyn *et al.,* 2018; Wang *et al.,* 2015). This platform was chosen over Zooniverse (the largest CS platform) or other project builders, for example, due to the range of functionality it provides. Most project builders are new in development, and, as such, the platform was selected based on its applicability to the project designed here (Liu *et al*., 2021). For example, Zooniverse is based on data collection through participants classifying images. This was not suitable for this project as the satellite images were classified by the model; it was the outputs of said model that required ground-truthing. Citsci.org provided functions to create online data entry forms, present and map the location of data entries (highly applicable to the aims of this survey), upload resources (e.g., guides, hard copy survey forms, model outputs), have discussion forums, and tabs to present analysis and results. Ctisci.org also has a mobile phone app that participants could access in field to view the project’s page.

6.3.2.5 Volunteer Recruitment and Training

Participants were recruited through collaboration with Butterfly Conservation and other connections. Information regarding the project and how to get involved was disseminated via organisational news outlets (newsletters/email communication/social media posts) (see Appendix E-4). The project website was also promoted on social media platforms, such as twitter and Instagram. The project was further promoted during outreach events, such as at the Glasgow Science Festival (and its Science on the Sofa) and Pint of Science (Appendix E-5). There was an attempt to instigate the project, both the CS survey and the RS aspects, into school curriculums (to benefit both environmental and educational goals). However, time restraints and the pandemic limited this plan to initial discussions. This is further discussed in chapter 6.

Butterfly Conservation volunteers typically have substantial experience with data collection and butterfly transects, for which they may receive training from Butterfly Conservation staff and other volunteers. Information is also available on the project platform online regarding identifying species, how to set up and conduct transects, and access to Butterfly Conservation's own training events (see Appendix E-6). The project platform had further ID guides available for both plant and butterfly species, guides on how to conduct the transects (handouts), apps that can be used whilst out surveying, and contact details of researchers and online forums to discuss any issues raised. Survey events were also set up on Eventbrite (2023) where I was available for outlining the methodology and helping with identification (see Appendix E-7).

6.3.2.6 Project Pilot and Method Refinement

The project was made live towards the end of July 2022 following ethical approval by the University of Glasgow, College of Social Sciences Ethics committee. The initialisation of the project finished in September 2022 when grass die out was occurring. During the 2022 season, a group was taken out with the me, on the Isle of Harris, to test the survey methods and to gain feedback. The in-field survey was successful, and these methods were left as designed. During the 2022 season, secondary data on previous sites of SRGs were provided by members of the public. These were then cross-referenced with thematic maps derived from remote sensing images. Issues of GPS accuracies and out-dated survey dates were brought to my attention. This determined that previously collected location data on SRGs could only be used if the sites were confirmed as an SRG type by the citizen scientist or data provider since 2021. Any GPS locations given had to be taken within a core habitat, or where this was not possible, details of the surrounding habitats had to be listed.

The project was reinstated for the beginning of the 2023 survey season (May-July) to increase data capture. Targeted project promotion occurred for the 2023 season, due to an initial low response in 2022, which involved museum talks, group invites, networking via known contacts, and further social media promotion. These were spread across Scotland to enlist people from as many areas as possible. I was available at these survey days to help explain the methodology and with plant identifications. For four survey days, I chose predicted areas from model predicted locations to survey due to accessibility requirements and ability to park, for example. Where survey days had been set up with organisations (Plantlife and The Conservation Volunteers), sites were often chosen by the organisation in areas they usually work in.

6.3.3 Sources of Additional Data

Previously confirmed areas of SRGs were supplied by a range of organisations including NatureScot (2023), the Botanical Society of Britain and Ireland (Miles, 2023), Plantlife (2022), and by individuals. To ensure these were up to date, only areas that were confirmed after 2020 were included for analysis. This data was provided either as polygon areas or point locations from specific coordinates. Where polygons were provided (such as from NatureScot) the data was filtered by date (post 2020) and SRG coverage. Only polygons with an SRG coverage of 100% could be included, as exact locations of SRGs within the polygons could not be provided. A negative 10 m (one pixel edge length) buffer was created against the polygon boundaries to remove edge effects from influencing surface reflectance of boundary habitats (see Figure 6-3 for model classification issues of small habitat patches). Random points were then generated within the buffer zone to identify single pixels within an SRG site. Raster values of the model predicted classification maps were extracted for each point within the buffer SRG zones to compare to the recorded SRG type.

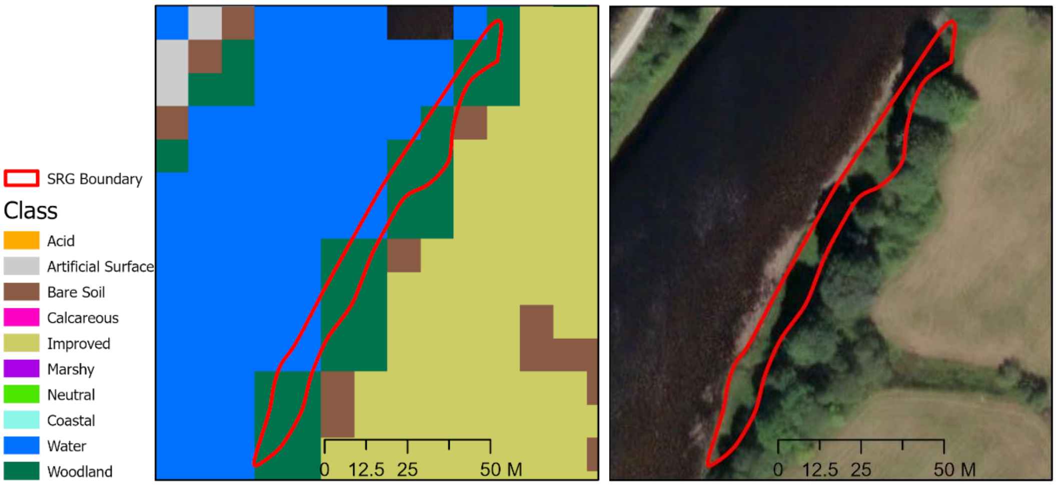


Figure 6-3. Issues of habitat edge effects from Sentinel-2's pixel size seen in some species-rich grassland sites.

6.3.4 Data Extraction and Analysis

6.3.4.1 Evaluating the Achievement of Open Science and Interdisciplinarity

In accordance with the aims and underlying principles of this thesis, the creation, implementation and reporting of this CS survey was evaluated for their apparent success and adherence to OS practices. Although there are few widely used approaches to achieve this, an evaluation framework, such as that by Kieslinger et al. (2017) may help projects assess their success of reaching project outcomes (for example, scientific, learning, or social). This framework is specific to CS surveys but could be adapted for this purpose of reflecting on the success of the interdisciplinary CS survey here and the associated thesis aims. In addition, my own framework for the determination of "openness" in biodiversity monitoring CS surveys, created in chapter 3, can be also utilised for this assessment (Suter *et al.*, 2023). As such, an evaluation table was adapted from both Kieslinger *et al*. (2017) and Suter *et al*. (2023) to investigate the strengths and weaknesses of the research here, questioning whether the engagement, interdisciplinarity, and open science outcomes were achieved (Table 6-4). Scientific goals were not discussed or evaluated in this framework (Table 6-4), as these are individually discussed in the analysis of each separate chapter (i.e. through answering each chapter's research questions), including this one: by investigating the success of mapping SRGs using the tool, and the effects participation had on this. The R code is found in Appendix E-8.

Table 6-4. Evaluation framework used in assessing the citizen science and research process of the thesis.

Outcome Focus	Process of Evaluation	Actors/Objects of Evaluation
Open Science	Quantitative scoring system based on OS practices implemented (Chapter 3).	Research design, implementation, data, and softwares
Interdisciplinarity	Qualitative observations – how and where were the disciplines combined.	Research design, implementation, and data products
Stakeholder Engagement	Quantitative descriptive statistics – number of stakeholders engaged, individual and repeat participation, outreach events. Qualitative and demographic information.	Participants and stakeholders

6.3.4.2 Evaluating the Success of Mapping Species-Rich Grasslands and *A. artaxerxes* habitat

As the data was categorical (SRG type), the SRG classes needed to be coded to values that matched the raster values associated with the SRG predictions of the classification maps. The habitat classes and their coded raster values can be seen in Table 6-5. The citizen ground-truthed SRG types were coded, and their locations points were used to extract the raster values (and corresponding SRG type) from the model predicted habitat classification maps. These values were used to investigate the SRG class agreement between citizens' ground-truthed field observations versus the model predicted locations.

Table 6-5. Habitat class included in a habitat classification model to predict species-rich grasslands, and their associated raster values.

Habitat Class	Associated Raster Value
Acid SRG	1
Artificial Surface	2
Bare Soil	3
Calcareous SRG	4
Improved Grassland	5
Neutral SRG	6
Marshy SRG	7
Coastal SRG	8
Water	9
Woodland	10

A defined scale was used to test the level of agreement between the model predicted and citizen ground-truthed SRG locations. The agreement level was determined for an exact location (that has both a model predicted pixel and a corresponding citizen recorded GPS point), rather than the total number of model predictions and citizen observations regardless of location. Full agreement was defined as a value of 1, which was given if the model predicted SRG class and the citizen observed SRG class matched exactly e.g., the model predicted a neutral SRG location and the citizen ground-truthed observation also found that location to be a neutral SRG. Partial agreement was defined as a value of 0.5, which was given if either the model-predicted SRG class or the citizen ground-truthed SRG class was one type of SRG class, whilst the other was a different SRG class e.g., model predicted acid SRG, but field observed calcareous SRG. No agreement was defined as a value of 0, which was given if one of the data sets (either the model prediction or the citizen ground-truth) suggested an SRG and the other did not e.g., model predicted marshy SRG but citizen ground-truthed woodland.

Chi squared goodness-of-fit tests (X^2) were used to investigate the agreement between RS data and CS data. If the model predicted total number of SRG locations and specific SRG classes well, there would be an equal number of model-predicted versus ground-truthed SRGs. Categorical variables were coded to compute the X^2 values. An initial binary association investigated the total number of model-predicted versus citizen ground-truthed SRGs. For this analysis, model-predicted SRG locations were coded to 1 and citizen ground-truthed SRG locations were coded to 2 to count the frequencies. Binary analyses for number of model-predicted versus citizen ground-truthed SRG locations were also done by SRG class. Each SRG class kept their original raster class number as seen in the model outputs (Table 6-5). For the binary analyses, 2x2 contingency tables were used whereby the group levels were the model-predicted and ground-truthing results and the expected values were calculated based on an equal split of the total number of recorded data points per class (both model and ground truthing frequencies) (Table 6-6).

Table 6-6. Set-up of Chi-square test for goodness-of-fit between the frequency of model-predicted versus citizen ground-truthed species-rich grassland classes.

Data Set	Actual Observed Number	Theoretical Expected Number
Model SRG	67	91.5
Ground-truthed SRG	116	91.5

Chi-squared tests for independence (or Fisher’s Exact test where more than 20% of the cells had values less than five) were used to explore whether there was an association with the agreement level between model predicted versus citizen ground-truthed SRG locations based on 1) SRG class, 2) data provider (e.g., which organisation the data was submitted by), 3) participant experience, and 4) participant confidence (Table 6-7). If there were differences in the level of agreement between data providers, participant experience and confidence were used to further explore these differences.

Table 6-7. Set-up of Chi-square test for independence between the frequency across agreement levels between model-predicted versus citizen ground-truthed species-rich grassland classes by plant identification experience (none, moderate, advanced).

Agreement between model predicted and citizen observed SRG type	Plant Identification Experience		
	None	*Moderate*	*Advanced*
No Agreement (0)	2	7	37
Partial Agreement (0.5)	0	3	29
Full Agreement (1)	1	3	17

It was not possible to assess the model success of discovery of *A. artaxerxes* habitat from the citizen science survey. No data on *A. artaxerxes* sightings or *H. nummularium* observations were submitted. There was also no citizen ground-truthed coastal SRG location data points submitted and, as such, this SRG class was not included in the analysis. Analysis was conducted in R (v 3.6.3) and ArcGIS Pro (v 3.0.3). Descriptive statistics were used to explore volunteer uptake, demographics, and habitat coverage.

6.4 Results

6.4.1 Evaluating Adherence to Open Science Practices, Interdisciplinarity, and Stakeholder Engagement

The use of OS practices within the CS survey was assessed using the scoring system created in chapter 3 (Table 6-8). Currently, the project has an open preregistration, data is openly available on the project platform, the open software R has been used, however, the licensed but transferable software ArcGIS was also utilised. The project is not finalised and is still in the process of being published, thus scoring of the data management plan, code, and access is not yet possible. However, with the planned open science practices outlined in Table 6-8, the final average openness score for Ecosystem Explorers is predicted to be 0.92.

Table 6-8. Scoring of Open Practices within the Ecosystem Explorers citizen science survey.

Preregistration	Data Management Plan	Data	Code	Software	Access
Open and available (1)	Open and to be published with thesis and any associated publications (1)	CS observations are openly available on the project platform. Subsequent publications to have linked data either via the publisher or UoG Enlighten (1)	Subsequent publications to have open code, published either on a repository such as Github or as supplementary materials linked to the published papers and the thesis (to be accessible on Enlighten)(1)	All analysis was conducted with R, an open software. Some data extraction was conducted with ArcGIS Pro, a licensed software, with transferable use to the open software QGIS (0.5)	Publications are/will be open access (see chapter 3 associated paper), and the CS results will be available on the project's open platform (1)

The RS outputs developed in previous chapters were effectively used both to design and inform the deployment of a CS survey for biodiversity monitoring, and to be evaluated by said survey. RS data was used in chapter 4 to create outputs (maps, location points, predictions) that were fed into the CS survey. On the project platform, 16 habitat prediction maps and 1033 predicted GPS points were available to participants. Participants were able to view habitat prediction maps and locations for their surveying decision making. Individual maps or locations were also provided to participants who directly requested specific areas to be mapped.

The project itself was able to engage stakeholders across government, NGOs, and local councils and networks. This collaborative work was conducted with: Butterfly Conservation, NatureScot, Plantlife, the Botanical Society of Britain and Ireland, the Glasgow and Clyde Green Network, The Conservation Volunteers, and the Perth and Kinross Countryside Trust. Outreach events were conducted at the Glasgow Science Festival (2021), The University of Glasgow Climate Action Day (2021) the Tayside Recorders' Day (2022), and Pint of Science (2023), along with associated newsletter and social media contributions. Individually, 14 participants were signed up to the project platform. Furthermore, 31 individuals (including representatives from related organisations) were involved in eight planned survey days, with a maximum of 10 on one day. Three individual participants completed more than one survey. Little demographic data outside participant numbers and age group were collected. The implications surrounding this are further found in the discussion.

6.4.2 Organisational, Spatial, and Demographic Reach of Submitted Survey Locations

There were 119 location points available for analysis after data cleaning to remove those sites with edge effects, sites that were neither model-predicted or ground-truth observed SRGs, or sites where no cloud free acquisitions could be obtained. The citizen science platform itself received 26 of these, reduced to 20 after data cleaning, from 14 individuals who signed up to the project platform. On organised survey days, participants split into groups rather than conducting individual transects due to participant preferences and the limitations of only having myself representing Ecosystem Explorers. NatureScot data consisted of 76 final location points (Scobie, 2023), Plantlife data consisted of 20 final data points (Jones, 2023), the BSBI consisted of three final data points (Miles, 2023).

The ground-truthed data covered a range of locations across Scotland (Figure 6-4). Most areas surveyed were in the Cairngorms National Park. Data is largely missing from the west coast, the highlands, and the coastal margins of the Borders, and Dumfries and Galloway.

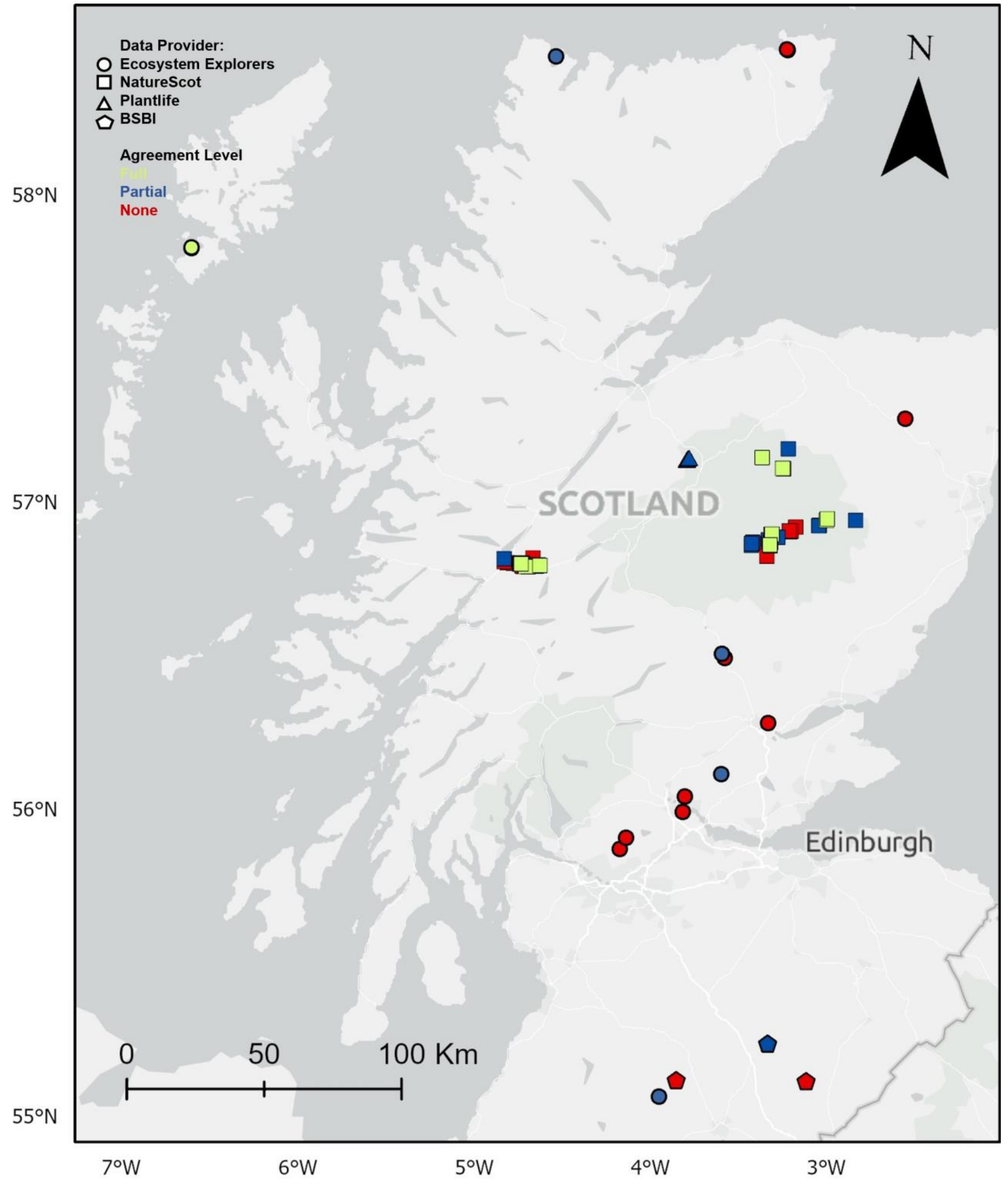


Figure 6-4. Ground-truthed survey locations across Scotland by organisation: Ecosystem Explorers (circle), Nature Scot (square), Plantlife (triangle), and BSBI (pentagon); and by level of agreement with the model predicted SRG class: full agreement (green), partial agreement (blue), and no agreement (red).

The participants ranged in age from under 18 to over 64. Where participants were under 18, they were surveying with consenting parents. Participant experience also ranged from no previous identification knowledge to advanced knowledge, and confidence in their data ranged from 1 (no confidence) to 5 (very confident) (Figure 6-5).

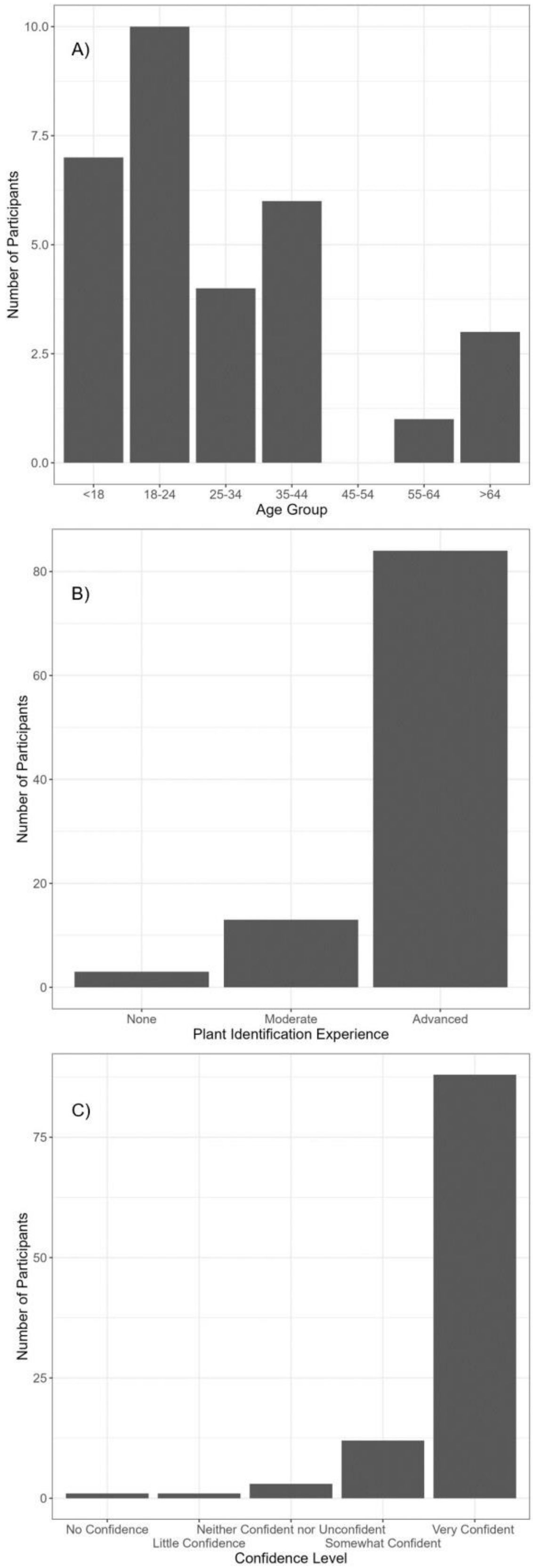


Figure 6-5. Number of participants in Ecosystem Explorers citizen science survey by A) age group, B) plant identification experience, and C) grassland classification confidence.

6.4.3 Total Agreement between Model Predicted, and Participant Observed Species-rich Grasslands, and Agreement by Species-rich Grassland Class

The data analysed the agreement between SRG locations that were ground-truthed in field by participants and predicted by the model. This included the model predicting SRGs where none were found and the model predicting other classes where SRGs were found. Of the 119 analysed locations, there was a 17.65% full agreement between model-predicted and citizen ground-truthed SRG locations. There was a 36.13% partial agreement, where the model predicted one type of SRG class but the citizens ground-truthed a different SRG class. There was no agreement with 46.22% of locations, however (Figure 6-6).

Of the 119 analysed locations, 67 of these were predicted as an SRG type by the model. Of these 67, 95.52% of these were confirmed as one type of SRG class, with 31.34% of these confirmed as the correct SRG class. Only 4.57% of these 67 were predicted an SRG where there were none. However, 116 of the 119 sites were confirmed SRGs by participants. The model misclassified 44.83% of these sites (not including the sites with partial agreement, as these were still classified as one SRG class), ruling them out as SRGs where their presence was confirmed. This suggests the model is underestimating areas of SRGs.

The null hypothesis states that there is no difference between theoretical expected versus citizen and model observed number of SRG locations between total model-predicted and citizen ground-truthed SRG locations. The results from the X^2 test for goodness-of-fit showed that there was a significant difference between the citizen and model observed versus theoretical expected values ($X^2 = 13.12$, df = 1, $p = 0.0003$), and that the distribution of data is not occurring randomly. This means that there is poor alignment between the model predicted and the citizen ground-truthed SRG locations.

The apparent efficacy of the model in predicting SRG classes (acid, calcareous, marshy, neutral, and coastal), was highly variable. There was a 3.5%, 0%, 11.54%, 21.3% full agreement between model predicted and citizen ground-truthed acid, calcareous, marshy, and neutral SRGs, respectively (Figure 6-6).

The null hypothesis states there is no difference in the proportion of agreement between model predicted and citizen ground-truthed SRGs by SRG class. The results from the Fisher’s exact test found a significant difference of agreement across the SRG classes ($p = 0.017$). Therefore, the null hypothesis is rejected, indicating that different SRG classes are predicted or citizen ground-truthed with variable success (Figure 6-6).

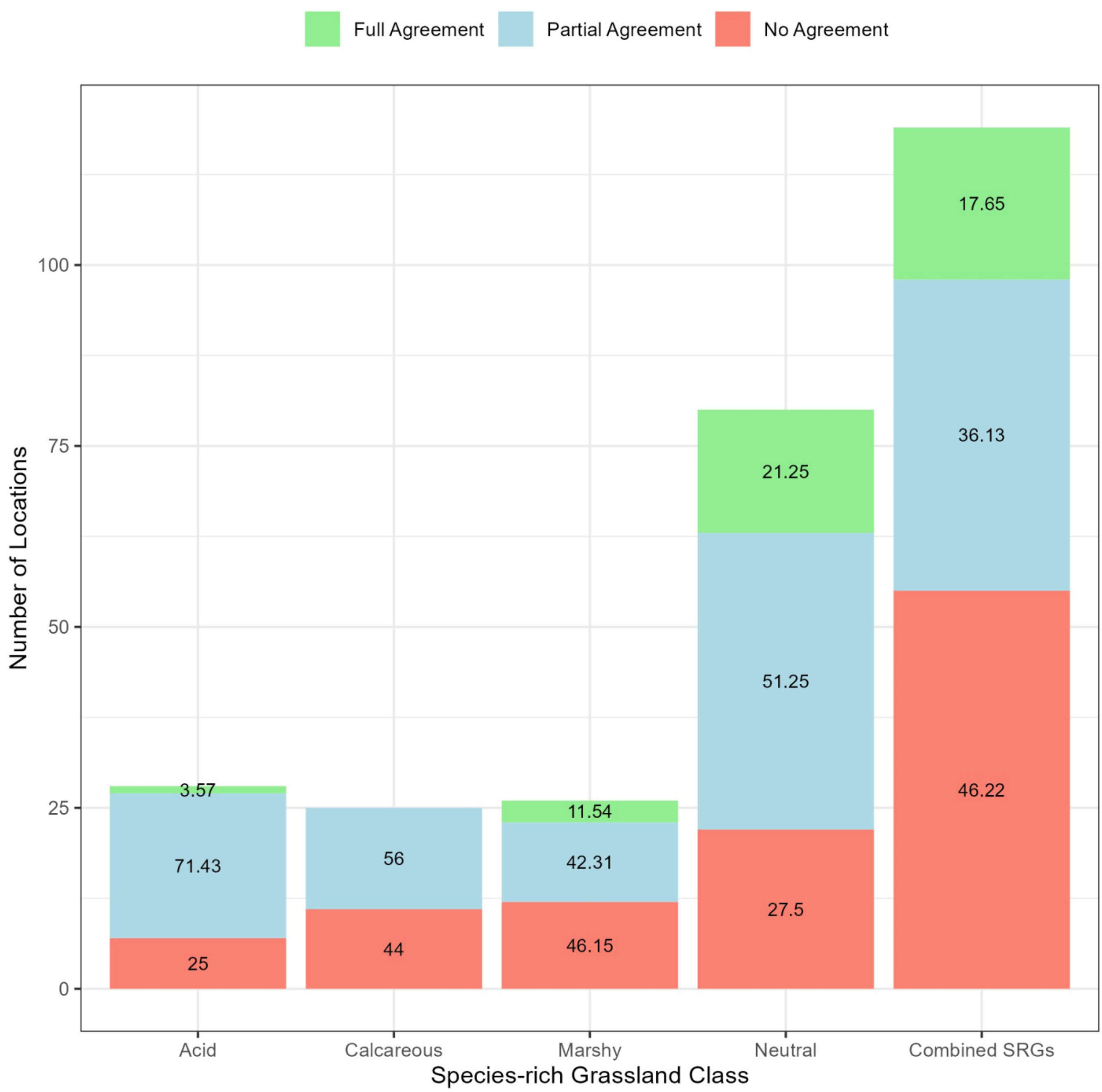


Figure 6-6. Agreement level between the number of model predicted and field observed SRG locations by SRG class (percentages provided in labels).

Of the analysed data, there were four locations that were predicted as acid SRG but 25 recorded field locations. There was one predicted calcareous SRG location but 24 recorded field locations. The model predicted three marshy SRG locations but there were 27 field locations. Finally, there were 59 predicted neutral SRG locations but only 38 recorded field locations. To further assess if there was a specific SRG class that was influencing the level of agreement between the model predictions and the citizens' ground-truthing, the X^2 goodness-of-fit tests were conducted. The X^2 for goodness-of-fit tests showed that there were significant differences between the model and citizen observed versus theoretical expected number of SRG locations, per each SRG class (Table 6-9). This suggests that the model predictions and citizens' ground-truthing are poorly aligned, per each SRG class.

Table 6-9. Results of the chi-squared goodness-of-fit tests investigating if the number of model predicted and citizen ground-truthed SRG locations differed from the theoretical expected numbers (if there was a high agreement between the model and the ground-truthing) by SRG class.

Class	Results	
	X^2	p
Acid	15.21	<0.0001
Calcareous	21.16	<0.0001
Marshy	18.24	<0.0001
Neutral	4.55	0.033

6.4.4 Agreement by Participant Group and Subsequent Effects of Experience and Confidence

Of the 76 locations submitted by NatureScot, there was a 22.36% full agreement between model predicted and citizen ground-truthed classifications. Of the 20 locations submitted through Ecosystem Explorers, there was a 20% full agreement between model predicted and citizen ground-truthed classifications. Of the 20 locations submitted by Plantlife and the three locations submitted by a BSBI member, there was a 0% full agreement between model predicted and citizen ground-truthed classifications.

The null hypothesis states that there is no difference between the participant groups as to whether their field observations fully agreed, partially agreed, or disagreed with the model predictions. The results of the Fisher's exact test show that there was no significant difference between the four participant groups ($p = 0.114$). Therefore, the null hypothesis is accepted (Figure 6-7).

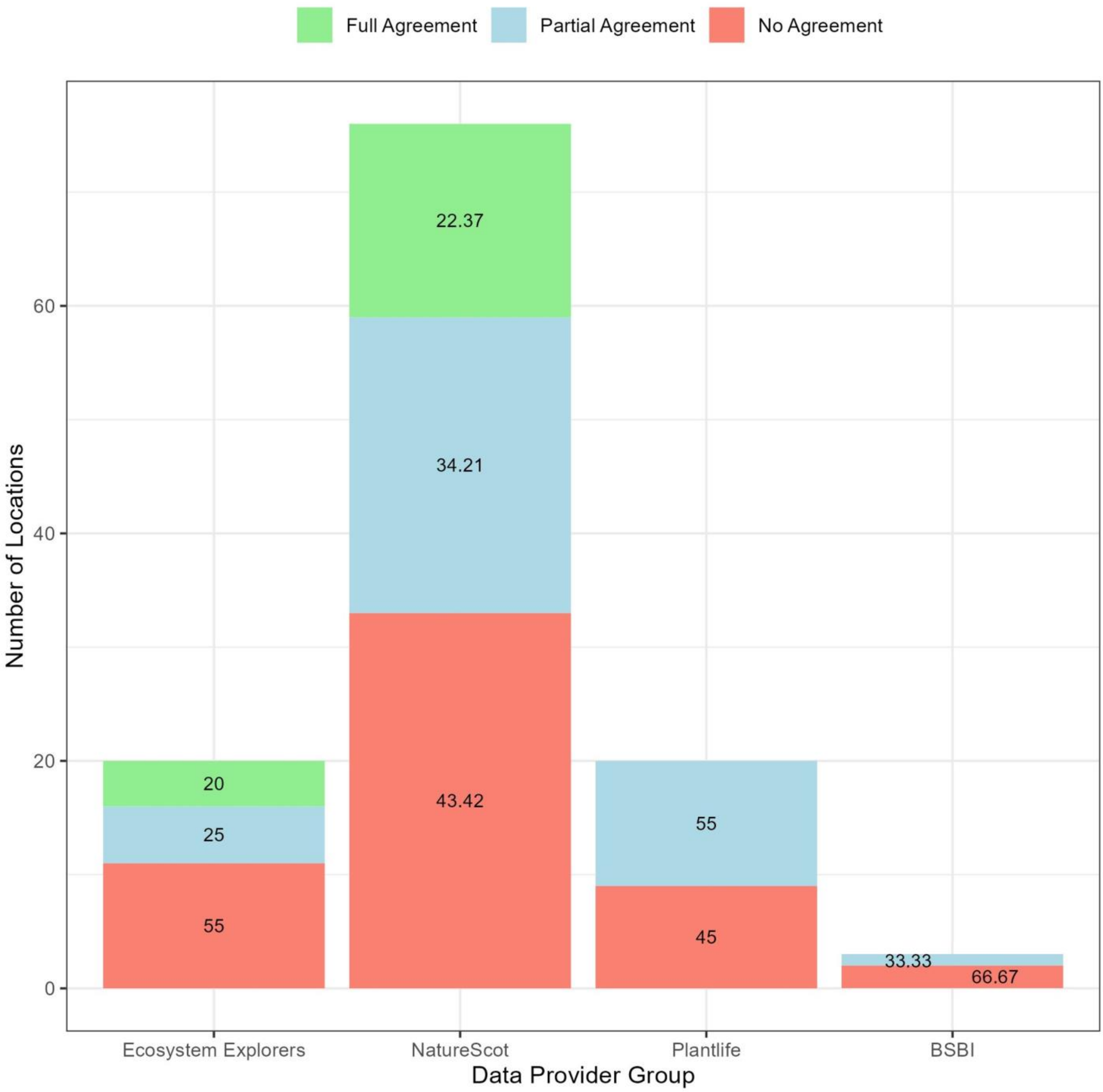


Figure 6-7. Agreement between the number of model predicted and field observed SRG locations by data provider group: Ecosystem Explorers, NatureScot, Plantlife, and BSBI) (percentages provided in labels).

Of the 119 final data points, 99 were used for the assessment of participant experience and confidence. This was the result of 20 surveys where the data provider could not assess experience and confidence. Most data were submitted by people with advanced experience and high confidence (83.84%). There was a range of participant experience and confidence from participants of Ecosystem Explorers. As there were no location points that had full agreement from advanced surveyors within Ecosystem Explorers, experienced surveyors are not thought to be skewing the data. Data from NatureScot was collected by professionals, hence the high number of advanced experience and high confidence. This data is still considered participatory citizen science or crowd sourcing through the process of data sharing.

The null hypothesis states that participant experience level would not affect whether the citizen ground-truthed locations fully agreed, partially agreed, or disagreed with the model predictions. The results of the Fisher's exact test on experience level show that there is no significant difference in participant plant identification experience level on the agreement level between citizen ground-truthed observations and model predictions of SRG locations ($p = 0.706$) (Figure 6-8).

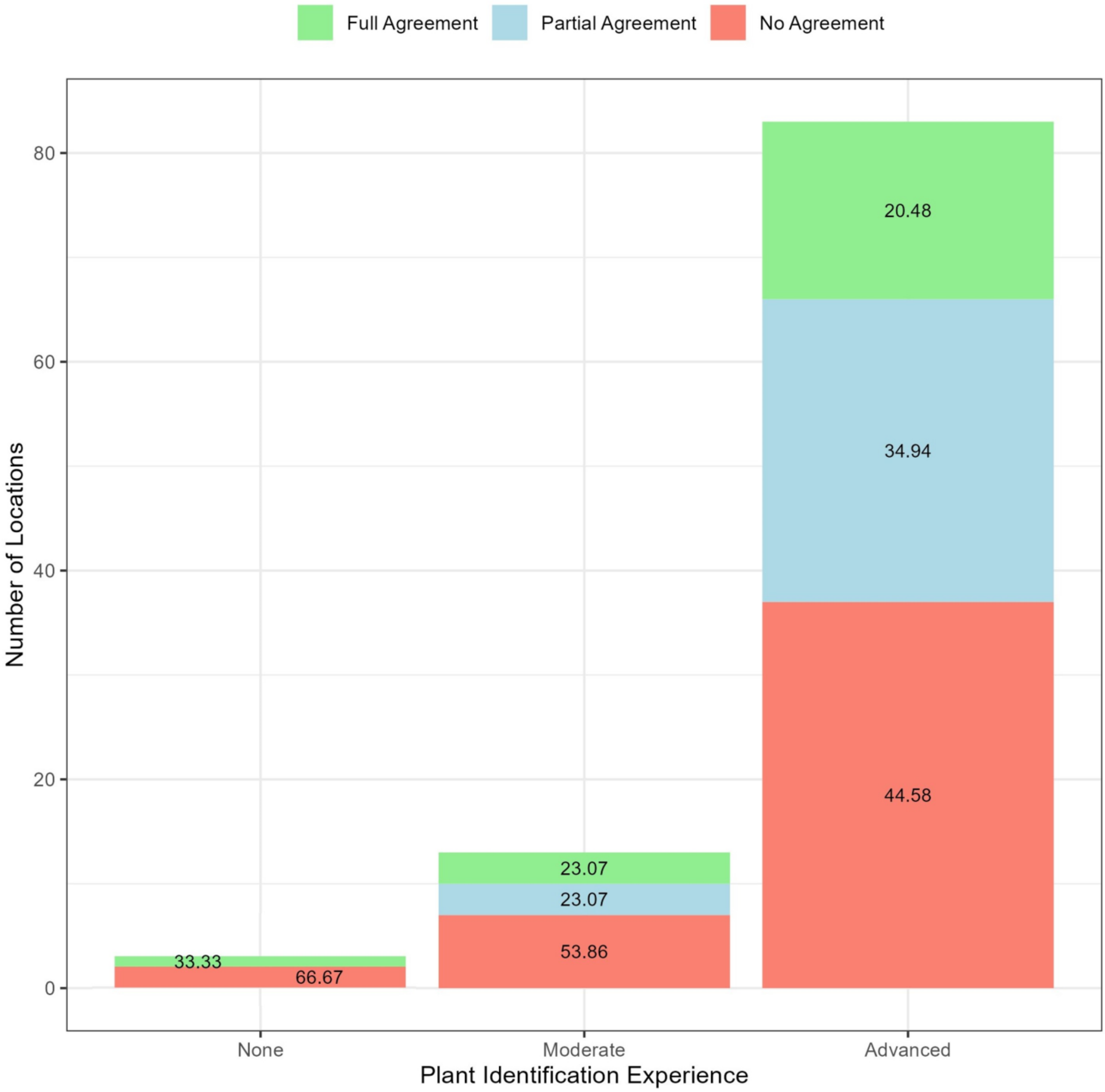


Figure 6-8. Agreement level between number of model predicted and participant field observed SRG locations by plant identification experience level (percentages provided in labels).

The null hypothesis states that participant confidence level would not affect whether the citizen ground-truthed observations fully agreed, partially agreed, or disagreed with the model predictions. The results of the Fisher's exact test on confidence level show that there is no significant difference in confidence level on the agreement level between citizen ground-truthed observations and model predictions of SRG locations ($p = 0.900$) (Figure 6-9).

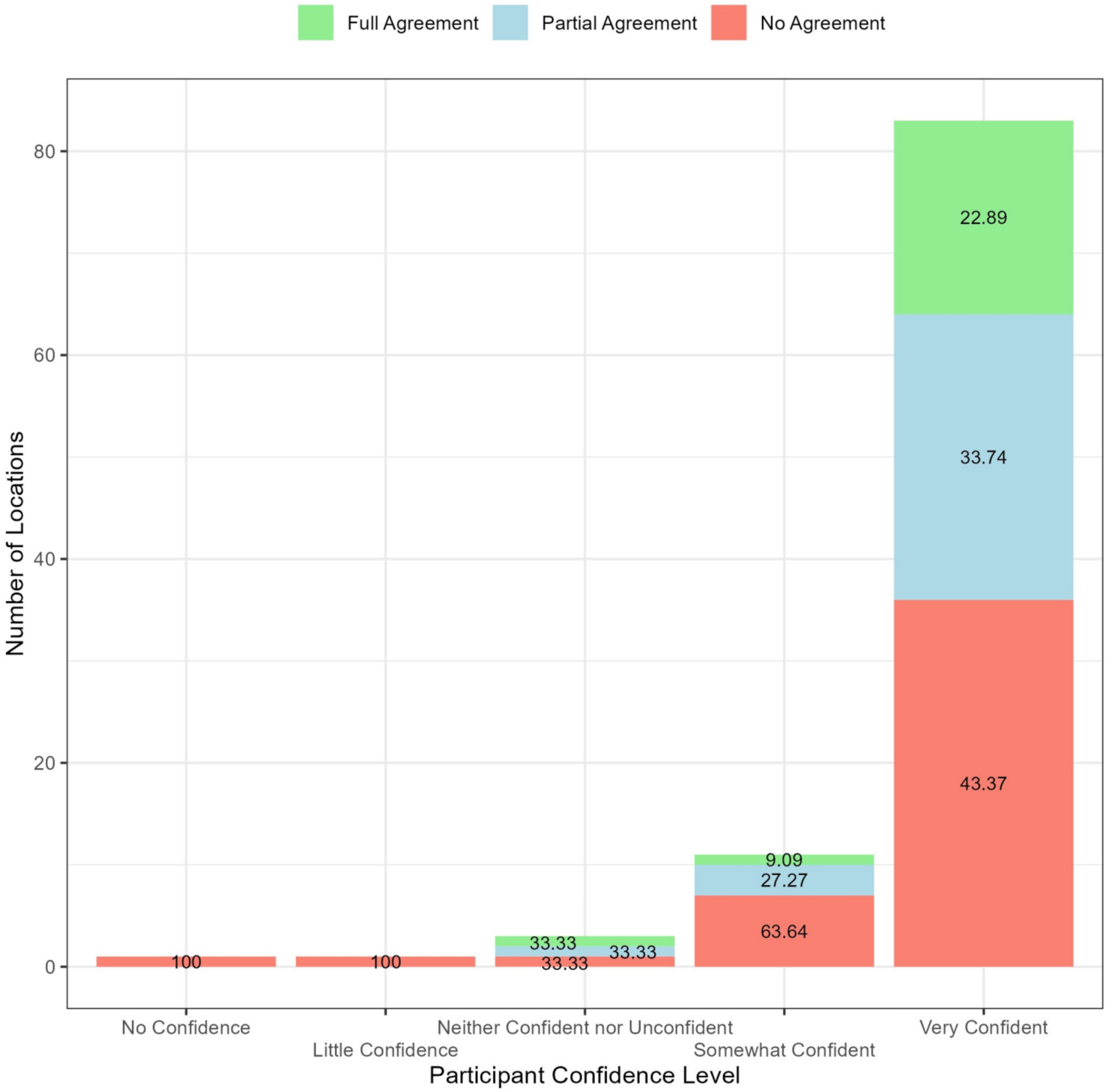


Figure 6-9. Agreement level between the number of model predicted and participant field observed SRG locations by habitat assessment confidence level (percentages provided in labels).

6.5 Discussion

The research enabled the creation of a habitat monitoring tool, Ecosystem Explorers, combining RS and CS methodologies. OS practices were able to be implemented throughout the research process, resulting in a predicted openness score of 0.92 when the project is finalised. Multiple stakeholders were engaged (from 7 organisations) throughout the research, forming important collaborations. Furthermore, 45 individuals participated in either solo or group surveys. New outputs for SRG mapping in Scotland were created from both RS SRG habitat prediction maps and citizen scientist observations, overall uniting the two disciplines.

The study found that there was a low agreement between model predicted and citizen ground-truthed SRG locations, which differed across SRG class but not by source of the ground-truthed data. Participant experience and confidence was also not found to have affected the level of agreement. Ecosystem Explorers is one of few projects that utilises RS data in a CS survey for biodiversity conservation. To our knowledge, no other studies utilise satellite imagery habitat predictions in a CS survey for mapping SRGs. As such, comparisons to previously published literature are difficult, especially as each CS survey works within unique contexts.

6.5.1 Utilising Open Science and Combining Citizen Science and Remote Sensing

The research here shows that it is possible to use OS practices in the creation of an interdisciplinary habitat monitoring tool, and perhaps facilitates this process. Certainly, through data sharing agreements collaborations become much more likely. This allowed further uptake of the tool and made reaching the scientific objectives of the research more feasible. Through these collaborative and OS efforts, other outcomes, not explicitly outlined along this research process, also become apparent and possible. For example, free survey days allowed participants access to learning and knowledge that they may not have otherwise engaged in. As such, it is likely that social outcomes were also achieved. However, this is not something that was analysed or taken into consideration throughout this thesis due to time constraints and researcher availability.

Few other RS CS projects exist within a biodiversity monitoring context, but where they do, they have also shown it is possible to utilise RS in CS surveys or vice versa. For example, Domingo-Marimon *et al.* (2022) and Purdy *et al.* (2023) successfully used S2 data and AVHRR/MODIS data (respectively) in CS projects to investigate tree phenology timings. The methodologies associated with these studies and ours, demonstrate the different ways in which RS and CS data can be combined: accessing previously collected CS data and acquiring RS data that aligns to the CS data collection period, or sending out volunteers to confirm the outputs generated by RS. Our study here essentially used both techniques; matching RS predictions to ground-truthed observations, submitted by participants that had previously surveyed SRGs versus ground-truthed observations collected by participants that had been directed to survey locations by the RS predictions.

Potentially, to further the engagement of citizens and their understanding of RS, there may be space for citizens to actively interact with RS products. For habitat monitoring, they may do this through creating their own classification maps for areas to survey or providing wider data on other habitats in the creation of their own prediction model, based on their local knowledge. This, however, would require greater training and resources related to the workflows associated with acquiring, processing, and analysing RS data.

6.5.2 Participation

Although a range of stakeholders and participants were enlisted in the project, there was unexpected low participation in the surveys, especially in its initial stages in 2022. There are several possible reasons for this. In the first season of the survey, participants were given large autonomy over where they could survey i.e., the decision was largely down to themselves. After communication with members of the public this was found to be too overwhelming for participants, which could be related to low confidence in themselves or difficulties interpreting the online maps and locations. When the survey ran in 2023, there was a targeted approach, specifically going to organisations and setting up defined survey days where I was available to provide training for surveying methodologies, help participants find a location they could access, and educate members on species identification. Other reasons could be related to the COVID-19 pandemic, as 2022 was the first year many people were going abroad again, CS saturation where there is too much choice of involvement, and difficulties associated with the methodologies and habitat knowledge.

It was noted that there is a preference amongst CS participants for biodiversity surveys that focus on individuals or groups of species rather than habitats (pers.comms., 2022). This vital information is important when setting up the design and goals of CS surveys, as ecosystem-based conservation clearly needs greater encouragement and involvement. Although there was a fauna surveying proponent to Ecosystem Explorers' methods, this did not seem to entice participants enough. It appears that there is no quantitative data on the number of species-based versus ecosystem/habitat-based CS surveys. CS surveys that monitor habitats/ecosystems do exist, however, these seem to be in largely aquatic environments (Huddart *et al.,* 2016; Schläppy *et al.,* 2017; Shuker *et al.,* 2017). Plant focused CS surveys tend to focus on groups (such as invasives), or specific species distribution or health data, and, as such, do not look at the entire habitat (Brown *et al.,* 2020; Dimson *et al.,* 2023; Marcenò *et al.,* 2021). This provides further evidence for the need of greater education surrounding ecosystems and the importance of these to provide habitat for the specific species that people may prefer to survey.

Not everyone who submitted data provided information on participant age, as it was not always possible to collect this retrospectively (i.e., from data that was provided by previously surveyed locations). Of age data that was submitted, it was found that most surveyors were between 18-24. There was a decline in surveyors after the age of 44, with only a slight increase after 64, despite Scotland having an ageing population (Scottish Government, 2021). Although a range of ages participated in the survey, it would be a target to get more people of middle-age and pensionable-age into CS. It was noted that participants were mostly not from ethnic minority groups. This demographic data was not specifically collected and aligns with Scotland's mostly white population (96% in 2011) (Scotland Census, 2011), but is noted as an area for much needed improvement. CS itself often fails to reach people of lower socio-economic status and ethnic minority groups, even in the Global North (Blake *et al*, 2020; Haklay, 2015; Mac Domhnaill *et al*, 2020). This further supports the need for projects, such as this, to be implemented into school curriculums. Due to the ecology of SRGs, they are usually found in natural areas away from built up urban spaces, especially for the context of this thesis. With most national parks only accessible from towns or city centres via car or public transport, this disadvantages ethnic minority groups the most, which were found to make up only 1% of national park visitors in the UK between 2005-2007 (Natural England, N.D.). Although a range of organisations were contacted to create survey days, the demographics of their own volunteers defined who was involved in the surveys. This can be amended by targeting specific organisations that engage ethnic minority groups, for example, the Black Environment Network (2023). Survey days can also be held in parks and recreation areas in cities to further this access to people with limited public transport. Equality in biodiversity monitoring is not only essential for representative sampling of data, but all people no matter their background, race, or socio-economic status should have access to these projects to increase scientific literacy and democratisation (Blake *et al.*, 2020).

6.5.3 Model Success

For the model to be successful, we would expect an equal number of model-predicted versus ground-truthed observed locations per SRG class and for the total number of SRG locations. The results suggest this is not the case. It is difficult to compare the success of habitat classification models based on S2 data in biodiverse grasslands such as the ones explored here. However, Merrington *et al.* (2021) found similar difficulties in a Scottish landscape with using S2 for high accuracy landcover classifications, even on train and test data. Similarly, work that is undergoing in Scotland using RS data to map SRGs is being done in the Cairngorms National Park in collaboration with NatureScot. However, these methods use aerial rather than satellite imagery because of spatial issues associated with satellites (pers.comms., 2023). S2 data was suitable to assess the phenological timings in the CS project by Domingo-Marimon *et al.* (2022), however, this was associated with tree species which are more equivalent to the spatial resolution of the RS sensor.

Earlier research in this thesis (chapter 3) showed high overall accuracies associated with mapping SRGs, similar to those reported in other studies mapping grassland vegetation (Pesaresi *et al.,* 2022; Rapinel *et al.,* 2019; Tarantino *et al.,* 2021). However, these studies did not look at the wide-scale applications of their models across unseen areas to further predict and map SRGs outside of the study sites. Le Dez *et al.* (2021) also noted the difficulties S2 experiences with picking up small habitat patches, which would be relevant to certain SRG areas across Scotland. It seems that although these models can accurately map an area associated with test and train data, the wider applications of these models may experience spatial limitations when considering the high variation associated with the SRG grassland classes used to train the model.

Although it could not be analysed, it is also assumed the model would not have been able to identify specific locations of *A. artaxerxes* habitat as the model predictions did not match well with the ground-truthed observations for calcareous grasslands. NBN Atlas occurrence data was considered for use in this analysis (which still complies as CS data). However, through investigating the submitted records, 14 observations were found across Scotland between 2021 and 2023 but abundances were not recorded. This would make it difficult to assess the actual model performance if an observation only consisted of one individual, for example, which are found on a centimetre scale and would not influence the spectral profile. Not only this, but there were also no observed SRG classes associated with the submissions, which would make the comparison between the model predicted classes not possible.

It would appear the 10 m pixels of S2 are too large to create an accurate habitat classification model for SRGs in reality. For example, when looking at small strips of SRGs or areas on habitat edges or boundaries, the habitat or land cover that makes up the greatest proportion of that pixel is going to determine the final classification based on the greatest influencing surface reflectance. Although this was accounted for in the methods, what could not be accounted for was the overlapping and mosaiced habitats of SRGs. It was observed *in situ* that different SRG classes can change from one half square metre quadrat to the next. It is also possible to find specific grassland communities of one class in the broader habitat of another grassland class. Therefore, capturing this variation at 10 m is nearly impossible. Highest accuracies are better guaranteed with both hyperspatial and multispectral data sets, such as that of the WorldView-2 satellite, that matches the spatial resolution of the habitats to be mapped (Merrington *et al.,* 2021). This makes providing an open access tool for grassland mapping particularly difficult when predictions rely on greater spatial resolution data to accurately map the quick successions of SRG communities. Open access satellite data is not available at the spatial resolution needed to do this.

The results also suggest that participant experience and confidence is unlikely to have affected the success of the model. This is also demonstrated by the results suggesting that the source of data did not influence the success either. This supports other studies that CS can provide high quality observations for the purposes of biodiversity monitoring, especially when adequate guides and training is made available (Jäckel *et al.,* 2021; Leocadio *et al.,* 2021; van der Velde *et al.,* 2017). It was noted that most people who participated in the survey did have some previous experience in plant identification. This could suggest that people who have a prior interest to botanical surveying or biodiversity conservation are more likely to participate in biodiversity CS surveys. As such, if the goals of the CS survey are to increase participant knowledge and engagement in biodiversity this may be an area that needs to be addressed. Unfortunately, the study here did not look at participant knowledge before and after participation in the survey, due to time restraints. However, it would be interesting to see the effects of participation on "bioliteracy", interest in biodiversity, and behaviour change resulting from participation, which could be an area for future research.

6.5.4 Limitations

Primarily the project utilised OS, however, there are still areas for improvement. For example, in the open scoring process, the use of open software was given a score of 0.5. The closed version of the GIS software ArcGIS was used for some data extraction, as opposed to the freely available QGIS. This was largely due to familiarity with the software platform and time restraints. A score of 0.5 is given as the applications used in ArcGIS are entirely transferable to QGIS and any associated instructions are freely available to apply to QGIS. It would still be encouraged to use open softwares where possible.

Time restrictions and low researcher numbers will have impacted participation levels (on top of reasons previously discussed), ultimately reducing stakeholder engagement. At survey days, having only me to work between surveying groups meant that fewer data points could be collected at these events, as participants had to be grouped together. Future efforts would target noticed pitfalls in stakeholder engagement and incorporate strategies and assessments to engage excluded societal groups for even greater scientific democracy. These limitations were largely unavoidable within the requirements of a PhD thesis. A longer running project, with greater researcher availability would serve to address these limitations.

The model was limited to the number of habitat/land cover classes that we could supervise it with. Due to time and labour restrictions, information could only be provided on collected SRG data and easily identified background classes from aerial imagery. Only 10 classes (five SRGs and five background classes were included in the model). There are many more habitat classes that could be assigned. The model, however, must assign each pixel in the image a classification from one of the 10 classes in the model. This means there is likely to be overestimation of the model classes where they have been assigned to areas of excluded classes. The most common model misclassification was that of improved grasslands (25.8% of classifications). This would suggest that the model is struggling to differentiate the unique features found within the various grassland classes, rather than from separate habitats altogether. From investigations it seems that sand is often classified as an artificial surface, shrubs (gorse and heather, for example) and hedgerows appear to be classified as woodland. Other excluded habitats include wetlands, which could have been classified as either marshy or water, moorland is likely to have been classified as acidic grassland, unimproved species-poor grassland could have been classified as improved grassland or one of the SRG classes. There is also a skew of SRG coverage in favour of neutral SRGs due to their dominance across Scotland, whilst calcareous grasslands are the rarest form in Scotland, and, as such, limited training data could be collected for this, and some of the other, SRG classes.

Bias was introduced to the CS methodology through the targeted survey approach in 2023, where survey locations were not able to be randomly chosen. This could not be avoided due to the need to find surveyable locations that were close to cities, had available parking or public transport, and were a model predicted SRG type to gain a representative sample. However, this is often the case in CS surveys where a specific species or habitat is needed to be monitored, and, as such, these areas are pinpointed. It was also not of interest to determine how well the model performed in general (e.g., classifying background classes). Only 4 of the 119 final locations were chosen by me.

6.5.5 Future Considerations

What we have learnt from this study is that RS outputs can be used in a CS survey to widen spatial reach of habitat mapping and biodiversity monitoring. However, the specific open habitat classification model was not successful at directing participants, and wider conservation organisations such as Butterfly Conservation to new areas of SRGs with high accuracy. This is somewhat surprising, given the high model accuracy (98.6%) found in chapter 3. Although some unidentified SRG locations may be located through this method, particularly neutral classes, it appears participants would more likely be led to areas of another habitat or unimproved/rough grassland that is not necessarily species-rich. This makes providing an OS framework where CS and RS can be combined in the context of SRGs near impossible. It may be that the use of open access S2 data is more applicable for predicting habitats with features of a larger spatial scale e.g., ancient woodlands, or even peatlands which are less diverse and less easily confounded by spectral similarities, being more distinct from other habitats. In the context of SRGs, higher spatial resolution data would be needed, which is more likely to come with associated costs.

This is problematic in a time where OS is of greater recognition and requirement for fulfilling targets, such as the SDGs. Needing access to closed RS data seems counterproductive to reversing the effects of biodiversity loss and climate change and makes it difficult for an open global biodiversity monitoring approach to be applicable across habitats, species, and regions if an OS tool does not work in all biodiversity contexts. We move for a call towards even further openness regarding RS and Earth observation data of all spectral and spatial resolutions, as there is high potential for conservation targets to be achieved through integration, collaboration, and scientific democratisation; the feasibility of which is just not currently in place.

6.5.6 A Methodology for Combining Remote Sensing and Citizen Science for Habitat Mapping

Although the specific S2 model did not perform well when extrapolated at a larger spatial scale (compared to the model’s initial success) at predicting areas of SRGs, we propose a methodology that combines RS and CS for future biodiversity mapping and monitoring resulting from this research and the research throughout this thesis (Figure 6-10):

1) Identify priority “biodiversity mapping unit” and invite collaborators and stakeholders: governments and organisations with policy and goals surrounding the biodiversity mapping unit;

members of the public where the conservation of the biodiversity mapping unit directly affects them.

2) Set up a collaborative or co-created citizen science survey with identified collaborators and stakeholders, holding meetings to determine the goals of the citizen science survey e.g., habitat mapping, biodiversity monitoring, public awareness of biodiversity loss, policy implementation.
3) Identify where remote sensing can be used in the survey and that data are fit-for-purpose surrounding the biodiversity mapping unit (considering spatial and spectral scale and choice of sensor). Where applicable this should be open source. However, where this is not possible, identify what data can be shared and how the research can be made as open as possible (open access software, code, results).
4) Identify how the public can participate in the citizen science survey and interact with the remote sensing data to successfully predict the biodiversity mapping unit e.g., ground-truthing model results (as seen here), providing location data on other habitat classes to inform the model training, collecting environmental data for creating habitat classification schemes.
5) Provide educational resources to stakeholders and participants e.g., identification guides and resources for the biodiversity mapping unit. However, if possible, also promote participant interaction with tangible remote sensing data e.g., visualisations of the satellite imagery; working with mapping software (free courses with QGIS) to map or identify their own locations; providing workflows of how maps are created.
6) Allow and promote continuous feedback and constant communication with stakeholders and participants for transparency and increased trustworthiness, whilst hopefully initiating higher success of the citizen science survey through updates and adaptations of methodologies – nothing should be fixed but a flexible approach should be adopted to the continued input of new information.

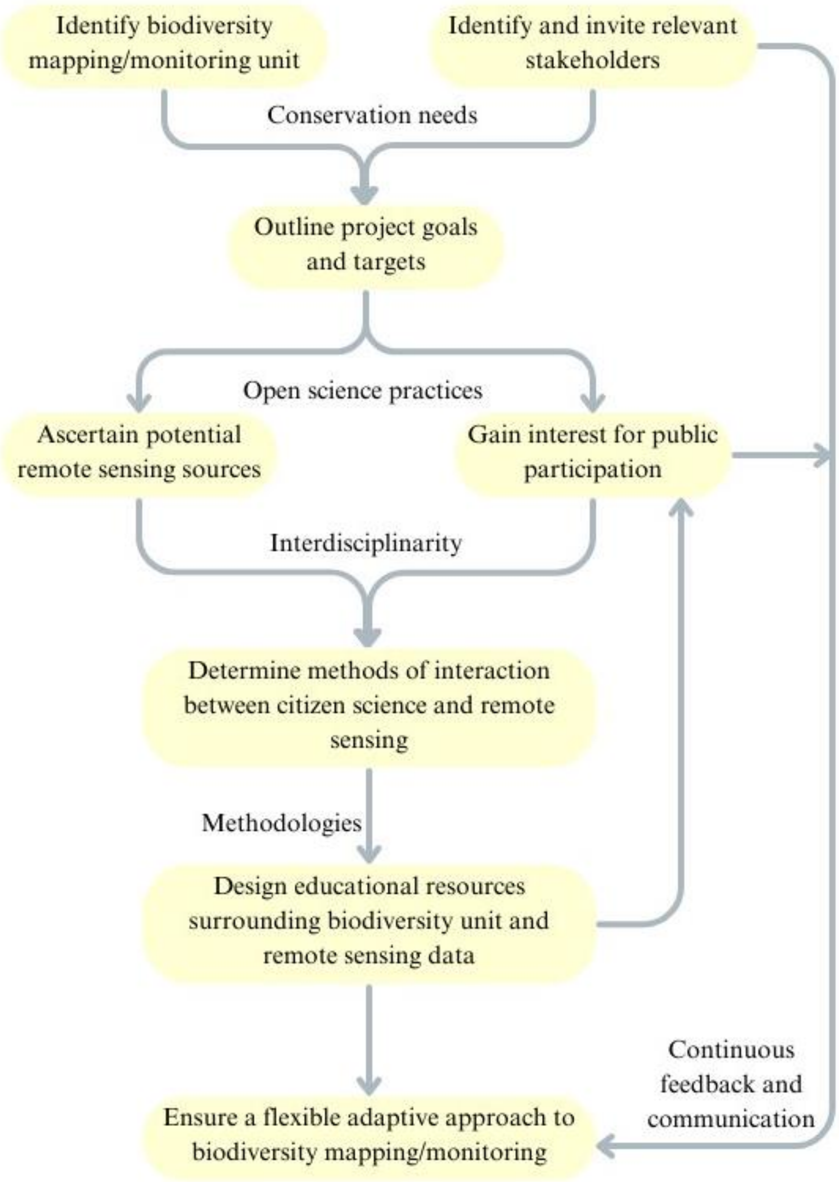


Figure 6-10. Workflow for designing a methodology that combines remote sensing and citizen science for future biodiversity mapping and monitoring.

6.6 Conclusion

Two commonly used approaches in biodiversity monitoring and land cover mapping are citizen science and remote sensing. However, the two approaches often never meet, to the disadvantage of biodiversity conservation. Both approaches can increase data collection capacity and widen spatial reach of monitoring and mapping attempts and together target the drawbacks of each tool. As such, this research aimed to combine citizen science and remote sensing in a tool for priority habitat mapping, specifically regarding species-rich grasslands, a degraded habitat, in an open science approach.

An accessible biodiversity monitoring tool was achieved through implementing open science practices and helped to cross discipline boundaries. Remote sensing outputs were able to be utilised in a citizen science survey that the public could interact with. The results show that there was a poor alignment between the open-source remote sensing data compared to citizens' ground-truthing for SRG mapping, either by overall locations or by SRG class. Participant experience and confidence, as well as where the data was sourced from did not affect how much the predicted model locations agreed with the citizen ground-truthed locations and, as such, the success of the model classifications. The results highlight the need for increased open access to high resolution remote sensing data to create an open access tool that utilises both remote sensing and citizen science for habitat mapping in all contexts. Although the specific remote sensing data was not preferable for predicting SRGs to send participants to survey, a methodology is designed for the ways in which these tools can be merged for mapping other priority habitats and biodiversity units.

References

Adler, F.R., Green, A.M. and Şekercioğlu, Ç.H., 2020. Citizen science in ecology: a place for humans in nature. *Annals of the New York Academy of Sciences*, *1469*(1), pp.52-64.

Anderson, C.B., 2018. Biodiversity monitoring, earth observations and the ecology of scale. *Ecology Letters*, *21*(10), pp.1572-1585.

Bardgett, R.D., Bullock, J.M., Lavorel, S., Manning, P., Schaffner, U., Ostle, N., Chomel, M., Durigan, G., L Fry, E., Johnson, D. and Lavallee, J.M., 2021. Combatting global grassland degradation. *Nature Reviews Earth & Environment*, *2*(10), pp.720-735.

Barrable, A., Booth, D., Adams, D. and Beauchamp, G., 2021. Enhancing nature connection and positive affect in children through mindful engagement with natural environments. *International Journal of Environmental Research and Public Health*, *18*(9), p.4785.

Blake, C., Rhanor, A. and Pajic, C., 2020. The Demographics of Citizen Science Participation and Its Implications for Data Quality and Environmental Justice. *Citizen Science: Theory and Practice*, *5*(1).

Boyd, D.S., Foody, G.M., Brown, C., Mazumdar, S., Marshall, H. and Wardlaw, J., 2022. Citizen science for Earth Observation (Citzens4EO): understanding current use in the UK. *International Journal of Remote Sensing*, *43*(8), pp.2965-2985.

Brooks, T.M., Butchart, S.H., Cox, N.A., Heath, M., Hilton-Taylor, C., Hoffmann, M., Kingston, N., Rodríguez, J.P., Stuart, S.N. and Smart, J., 2015. Harnessing biodiversity and conservation knowledge products to track the Aichi Targets and Sustainable Development Goals. *Biodiversity*, *16*(2-3), pp.157-174.

Brown, N., Pérez-Sierra, A., Crow, P. and Parnell, S., 2020. The role of passive surveillance and citizen science in plant health. *CABI Agriculture and Bioscience*, *1*, pp.1-16.

Carbutt, C., Henwood, W.D. and Gilfedder, L.A., 2017. Global plight of native temperate grasslands: going, going, gone?. *Biodiversity and Conservation*, *26*(12), pp.2911-2932.

De Meyer, K., Coren, E., McCaffrey, M. and Slean, C., 2020. Transforming the stories we tell about climate change: from 'issue' to 'action'. *Environmental Research Letters*, *16*(1), p.015002.

De Prins, J., 2022. Two urgent topics: Climate change and biodiversity loss. *Phegea*, *50*(1), pp.2-3.

De Sherbinin, A., Bowser, A., Chuang, T.R., Cooper, C., Danielsen, F., Edmunds, R., Elias, P., Faustman, E., Hultquist, C., Mondardini, R. and Popescu, I., 2021. The critical importance of citizen science data. *Frontiers in Climate*, *3*, p.20.

Dimson, M., Berio Fortini, L., Tingley, M.W. and Gillespie, T.W., 2023. Citizen science can complement professional invasive plant surveys and improve estimates of suitable habitat. *Diversity and Distributions*.

Domingo-Marimon, C., Masó, J., Cristóbal, J., Batalla, M. and Ninyerola, M., 2020, August. Remote sensing as a driving tool for Citizen Science phenology monitoring campaigns. In *Eighth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2020)* (Vol. 11524, pp. 596-605). SPIE.

Domingo-Marimon, C., Masó, J., Prat, E., Zabala, A., Serral, I., Batalla, M., Ninyerola, M. And Cristóbal, J., 2022. Aligning citizen science and remote sensing phenology observations to characterize climate change impact on vegetation. *Environmental Research Letters*, *17*(8), p.085007.

Dunkley, R.A., 2018. Space-timeScapes as ecopedagogy. *The Journal of Environmental Education*, *49*(2), pp.117-129.

ECSA (European Citizen Science Association). 2015. Ten Principles of Citizen Science. Berlin. http://doi.org/10.17605/OSF.IO/XPR2N

Fraisl, D., Campbell, J., See, L., when, U., Wardlaw, J., Gold, M., Moorthy, I., Arias, R., Piera, J., Oliver, J.L. and Masó, J., 2020. Mapping citizen science contributions to the UN sustainable development goals. *Sustainability Science*, *15*(6), pp.1735-1751.

Fritz, S., See, L., Carlson, T., Haklay, M.M., Oliver, J.L., Fraisl, D., Mondardini, R., Brocklehurst, M., Shanley, L.A., Schade, S. whenWehn, U., 2019. Citizen science and the United Nations sustainable development goals. *Nature Sustainability*, *2*(10), pp.922-930.

Gifford, R., 2011. The dragons of inaction: psychological barriers that limit climate change mitigation and adaptation. *American Psychologist*, *66*(4), p.290.

Golumbic, Y.N., Baram-Tsabari, A. and Koichu, B., 2020. Engagement and communication features of scientifically successful citizen science projects. *Environmental Communication*, *14*(4), pp.465-480.

Greving, H., Bruckermann, T., Schumann, A., Straka, T.M., Lewanzik, D., Voigt-Heucke, S.L., Marggraf, L., Lorenz, J., Brandt, M., Voigt, C.C. and Harms, U., 2022. Improving attitudes and knowledge in a citizen science project about urban bat ecology. *Ecology and Society: E&S 27 (2022), Nr. 2*, *27*(2), p.24.

Gunnell, J.L., Golumbic, Y.N., Hayes, T. and Cooper, M., 2021. Co-created citizen science: challenging cultures and practice in scientific research. *Journal of Science Communication*, *20*(5), pp.1-17.

Haklay, M., 2015. *Citizen science and policy: A European perspective*. Washington, DC: Woodrow Wilson International Center for Scholars.

Haklay, M.M., Dörler, D., Heigl, F., Manzoni, M., Hecker, S. and Vohland, K., 2021. What is citizen science? The challenges of definition. *The Science of Citizen Science*, *13*.

Hecker, S., Haklay, M., Bowser, A., Makuch, Z., Vogel, J. and Bonn, A., 2018. Innovation in open science, society and policy–setting the agenda for citizen science. *Citizen Science: Innovation in Open Science, Society and Policy*, pp.1-23.

Huddart, J.E., Thompson, M.S., Woodward, G. and Brooks, S.J., 2016. Citizen science: from detecting pollution to evaluating ecological restoration. *Wiley Interdisciplinary Reviews: Water*, *3*(3), pp.287-300.

Jäckel, D., Mortega, K.G., Sturm, U., Brockmeyer, U., Khorramshahi, O. and Voigt-Heucke, S.L., 2021. Opportunities and limitations: A comparative analysis of citizen science and expert recordings for 176ioacoustics research. *Plos One*, *16*(6), p.e0253763.

Johnston, A., Matechou, E. and Dennis, E., 2022. Outstanding challenges and future directions for biodiversity monitoring using citizen science data. *Methods in Ecology and Evolution*.

Kosmala, M., Crall, A., Cheng, R., Hufkens, K., Henderson, S. and Richardson, A.D., 2016. Season Spotter: Using citizen science to validate and scale plant phenology from near-surface remote sensing. *Remote Sensing*, *8*(9), p.726.

Kühl, H.S., Bowler, D.E., Bösch, L., Bruelheide, H., Dauber, J., Eichenberg, D., Eisenhauer, N., Fernández, N., Guerra, C.A., Henle, K. and Herbinger, I., 2020. Effective biodiversity monitoring needs a culture of integration. *One Earth*, *3*(4), pp.462-474.

Lambers, K., Verschoof-van der Vaart, W.B. and Bourgeois, Q.P., 2019. Integrating remote sensing, machine learning, and citizen science in Dutch archaeological prospection. *Remote Sensing*, *11*(7), p.794.

Le Dez, M., Robin, M. and Launeau, P., 2021. Contribution of Sentinel-2 satellite images for habitat mapping of the Natura 2000 site 'Estuaire de la Loire'(France). *Remote Sensing Applications: Society and Environment*, *24*, p.100637.

Lee, K.A., Lee, J.R. and Bell, P., 2020. A review of Citizen Science within the Earth Sciences: potential benefits and obstacles. *Proceedings of the Geologists' Association*, *131*(6), pp.605-617.

Leocadio, J.N., Ghilardi-Lopes, N.P., Koffler, S., Barbiéri, C., Francoy, T.M., Albertini, B. and Saraiva, A.M., 2021. Data Reliability in a Citizen Science Protocol for Monitoring Stingless Bees Flight Activity. *Insects*, *12*(9), p.766.

Liu, H.Y., Dörler, D., Heigl, F. and Grossberndt, S., 2021. Citizen science platforms. *The Science of Citizen Science*, *22*, pp.439-459.*;Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., Samson, R., Wagenknecht, K., Eds*, pp.439-459.

Lynn, S.J., Kaplan, N., Newman, S., Scarpino, R. and Newman, G., 2019. Designing a platform for ethical citizen science: A case study of CitSci. Org. *Citizen science*, *4*(1).

Mac Domhnaill, C., Lyons, S. and Nolan, A., 2020. The citizens in citizen science: demographic, socioeconomic, and health characteristics of biodiversity recorders in Ireland. *Citizen Science: Theory and Practice*, *5*(1).

Marcenò, C., Padullés Cubino, J., Chytrý, M., Genduso, E., Salemi, D., La Rosa, A., Gristina, A.S., Agrillo, E., Bonari, G., Giusso del Galdo, G. And Ilardi, V., 2021. Facebook groups as citizen science tools for plant species monitoring. *Journal of Applied Ecology*, *58*(10), pp.2018-2028.

Mazumdar, S., Wrigley, S. and Ciravegna, F., 2017. Citizen science and crowdsourcing for earth observations: An analysis of stakeholder opinions on the present and future. *Remote Sensing*, *9*(1), p.87.

Merrington, A.T., Milodowski, D.T. and Williams, M., 2021, October. Optimising remotely sensed land cover classification for habitat mapping in complex Scottish upland landscapes. In *Space, Satellites, and Sustainability II* (Vol. 11888, pp. 65-77). SPIE.

Miles, C., 2023. Data supplied as a Botanical Society of Britain and Ireland recorder.

Natural England., 2020. Climate Change Adaptation Manual (NE751). *[Online]*. Accessed 31st March 2022. Available at: http://publications.naturalengland.org.uk/publication/5679197848862720

Natural England., N.D. The Mosaic Model Access to Nature: Early Findings Paper. *[Online]*. Accessed 14th August 2023. Available at: https://publications.naturalengland.org.uk/file/6657310910840832

NatureScot., 2023. NatureScot (Scottish Natural Heritage), licensed under the Open Government Licence. Datasets: (i) CNP Grassland Mapping Project 2020 to 2022 (ii) Brae Lochaber Grassland Mapping 2020 to 2021.

Parrish, J.K., Burgess, H., Weltzin, J.F., Fortson, L., Wiggins, A. and Simmons, B., 2018. Exposing the science in citizen science: fitness to purpose and intentional design. *Integrative and Comparative Biology*, *58*(1), pp.150-160.

Pesaresi, S., Mancini, A., Quattrini, G. And Casavecchia, S., 2022. Functional analysis for habitat mapping in a special area of conservation using sentinel-2 time-series data. *Remote Sensing*, *14*(5), p.1179.

Plantlife, 2022. Plantlife Cairngorms Rare Plants and Wild Connections Project. [Data].

Price, E., Maguire, S., Firth, C., Lumber, R., Richardson, M. and Young, R., 2022. Factors associated with nature connectedness in school-aged children. *Current Research in Ecological and Social Psychology*, *3*, p.100037.

Purdy, L.M., Sang, Z., Beaubien, E. and Hamann, A., 2023. Validating remotely sensed land surface phenology with leaf out records from a citizen science network. *International Journal of Applied Earth Observation and Geoinformation*, *116*, p.103148.

Purnomo, E.P., Khairina, E., Fadhlurrohman, M.I., Andoko, B.W. and Priyanto, A., 2020. Managing biodiversity in dealing with sustainable development goals. *Mimb./Sos. Pembang*, *36*, pp.371-382.

Rahbek, C., 2011. The biodiversity crisis: Worse than climate change. University of Copenhagen. Available at: https://news.ku.dk/all_news/older/biodiversity/

Rapinel, S., Mony, C., Lecoq, L., Clément, B., Thomas, A. and Hubert-Moy, L., 2019. Evaluation of Sentinel-2 time-series for mapping floodplain grassland plant communities. *Remote Sensing of Environment*, *223*, pp.115-129.

Robinson, L.D., Cawthray, J.L., West, S.E., Bonn, A. and Ansine, J., 2018. Ten principles of citizen science. In *Citizen science: Innovation in open science, society and policy* (pp. 27-40). UCL Press.

Roe, D., Seddon, N. and Elliott, J., 2019. Biodiversity loss is a development issue. *International Institute for Environment and Development. Available at: https://pubs. Iied. Org/pdfs/17636IIED. Pdf (accessed 22 February 2021)*.

Sachs, J.D., 2012. From millennium development goals to sustainable development goals. *The Lancet*, *379*(9832), pp.2206-2211.

Saunders, M.E., Janes, J.K. and O'Hanlon, J.C., 2020. Moving on from the insect apocalypse narrative: engaging with evidence-based insect conservation. *BioScience*, *70*(1), pp.80-89.

Scannell, L. and Gifford, R., 2013. Personally relevant climate change: The role of place attachment and local versus global message framing in engagement. *Environment and Behavior*, *45*(1), pp.60-85.

Schläppy, M.L., Loder, J., Salmond, J., Lea, A., Dean, A.J. and Roelfsema, C.M., 2017. Making waves: marine citizen science for impact. *Frontiers in Marine Science*, *4*, p.146.

Schulte to Bühne, H. and Pettorelli, N., 2018. Better together: Integrating and fusing multispectral and radar satellite imagery to inform biodiversity monitoring, ecological research and conservation science. *Methods in Ecology and Evolution*, *9*(4), pp.849-865.

Scotland Census., 2011. Scotland's Census – general report. Accessed 22nd September 2023. Available at: https://www.scotlandscensus.gov.uk/media/jx2lz54n/scotland-s_census_2011_general_report.pdf

Scottish Government., 2021. A Scotland for the future: opportunities and challenges of Scotland's changing population. Accessed 22nd September 2023. Available at: https://www.gov.scot/publications/scotland-future-opportunities-challenges-scotlands-changing-population/pages/6/

See, L., Bayas, J.C.L., Lesiv, M., Schepaschenko, D., Danylo, O., McCallum, I., Dürauer, M., Georgieva, I., Domian, D., Fraisl, D. and Hager, G., 2022. Lessons learned in developing reference data sets with the contribution of citizens: the Geo-Wiki experience. *Environmental Research Letters*, *17*(6), p.065003.

See, L., Fritz, S., Dias, E., Hendriks, E., Mijling, B., Snik, F., Stammes, P., Vescovi, F.D., Zeug, G., Mathieu, P.P. and Desnos, Y.L., 2016. Supporting Earth-Observation Calibration and Validation: A new generation of tools for crowdsourcing and citizen science. *IEEE Geoscience and Remote Sensing Magazine*, *4*(3), pp.38-50.

Senabre Hidalgo, E., Perelló, J., Becker, F., Bonhoure, I., Legris, M. and Cigarini, A., 2021. Participation and co-creation in citizen science. *Chapter 11. In: Vohland K. et al.(Eds). 2021. The Science of Citizen Science.* Springer. https://doi. Org/10.1007/978-3-030-58278-4. Pp: 199-218.

Shuker, L.J., Gurnell, A.M., Wharton, G., Gurnell, D.J., England, J., Finn Leeming, B.F. and Beach, E., 2017. MoRPh: a citizen science tool for monitoring and appraising physical habitat changes in rivers. *Water and Environment Journal*, *31*(3), pp.418-424.

Singh, J.S., 2002. The biodiversity crisis: a multifaceted review. *Current Science*, pp.638-647.

Steffen, P.J.C., 2003. How long have we been in the Anthropocene era?. *Climatic Change*, *61*(3), p.251.

Stephenson, P.J., 2019. Integrating remote sensing into wildlife monitoring for conservation. *Environmental Conservation*, *46*(3), pp.181-183.

Stephenson, P.J., 2020. Technological advances in biodiversity monitoring: applicability, opportunities and challenges. *Current Opinion in Environmental Sustainability*, *45*, pp.36-41.

Stroud, S., Fennell, M., Mitchley, J., Lydon, S., Peacock, J. and Bacon, K.L., 2022. The botanical education extinction and the fall of plant awareness. *Ecology and Evolution*, *12*(7), p.e9019.

Tan, Y.R., Agrawal, A., Matsoso, M.P., Katz, R., Davis, S.L., Winkler, A.S., Huber, A., Joshi, A., El-Mohandes, A., Mellado, B. and Mubaira, C.A., 2022. A call for citizen science in pandemic preparedness and response: beyond data collection. *BMJ Global Health*, *7*(6), p.e009389.

Tarantino, C., Forte, L., Blonda, P., Vicario, S., Tomaselli, V., Beierkuhnlein, C. and Adamo, M., 2021. Intra-annual sentinel-2 time-series supporting grassland habitat discrimination. *Remote Sensing*, *13*(2), p.277.

The Black Environment Network., 2023. https://ben-network.org.uk/

Theobald, E.J., Ettinger, A.K., Burgess, H.K., DeBey, L.B., Schmidt, N.R., Froehlich, H.E., Wagner, C., HilleRisLambers, J., Tewksbury, J., Harsch, M.A. and Parrish, J.K., 2015. Global change and local solutions: Tapping the unrealized potential of citizen science for biodiversity research. *Biological Conservation*, *181*, pp.236-244.

Thomas, H., Ougham, H. and Sanders, D., 2022. Plant blindness and sustainability. *International Journal of Sustainability in Higher Education*, *23*(1), pp.41-57.

Turak, E., Harrison, I., Dudgeon, D., Abell, R., Bush, A., Darwall, W., Finlayson, C.M., Ferrier, S., Freyhof, J., Hermoso, V. and Juffe-Bignoli, D., 2017. Essential Biodiversity Variables for measuring change in global freshwater biodiversity. *Biological Conservation*, *213*, pp.272-279.

Van der Velde, T., Milton, D.A., Lawson, T.J., Wilcox, C., Lansdell, M., Davis, G., Perkins, G. and Hardesty, B.D., 2017. Comparison of marine debris data collected by researchers and citizen scientists: Is citizen science data worth the effort?. *Biological Conservation*, *208*, pp.127-138.

Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., Samson, R. And Wagenknecht, K., 2021. The Science of Citizen Science Evolves. *Chapter 1. In Vohland, K. et al (Eds.).(2021) The Science of Citizen Science*. Springer. https://doi. Org/10.1007/978-3-030-58278-4. Pp 1-12.

Wang Y, Kaplan N, Newman G, Scarpino R (2015) Correction: CitSci.org: A New Model for Managing, Documenting, and Sharing Citizen Science Data. *PloS Biology, 13*(12): e1002343. https://doi.org/10.1371/journal.pbio.1002343

Western, D., 1992. The biodiversity crisis: a challenge for biology. *Oikos*, pp.29-38.

Wetzel, F.T., Bingham, H.C., Groom, Q., Haase, P., Kõljalg, U., Kuhlmann, M., Martin, C.S., Penev, L., Robertson, T., Saarenmaa, H. and Schmeller, D.S., 2018. Unlocking biodiversity data: Prioritization and filling the gaps in biodiversity observation data in Europe. *Biological Bonservation*, *221*, pp.78-85.

Wilson, R.J. and Fox, R., 2021. Insect responses to global change offer signposts for biodiversity and conservation. *Ecological Entomology*, *46*(4), pp.699-717.

Xu, H., Cao, Y., Yu, D., Cao, M., He, Y., Gill, M. and Pereira, H.M., 2021. Ensuring effective implementation of the post-2020 global biodiversity targets. *Nature Ecology & Evolution*, *5*(4), pp.411-418.

Zawadzki, S.J., Steg, L. And Bouman, T., 2020. Meta-analytic evidence for a robust and positive association between individuals' pro-environmental behaviors and their subjective wellbeing. *Environmental Research Letters*, *15*(12), p.123007.

Chapter 7. Research Synthesis and Conclusion

7.1 Background

7.1.1 Anthropogenic Influences and Priorities

We live in an era that is defined by climate change and habitat loss, driven by anthropogenic activities that threaten the future of the planet (Leadley *et al.,* 2022). Awareness of these issues is not new, with global warming linked to fossil fuel use as early as the 19th century (Rathi, 2016). Early evidence of the onset and advancement of the biodiversity crisis includes: the impact of early hominid species' migration affecting the evolution of size in mammals, the effects of 17th century fishing and deforestation, and the 19th century effects of agriculture and rapid human population increase (Garthwaite, 2018; Weston, 2022). Despite both rising public concern, which has only grown since the 1980s, and a mounting body of evidence, globally we have been unsuccessful in responding to these problems, as seen in the continuous average decline in biodiversity (van Goethem and van Zanden, 2021).

Focuses on capitalism continue to exacerbate these problems. For example, economic growth has been embedded in policy priorities in the UK, since the 1980s. However, alternative viewpoints and models exist, for example, the Eco-Marxist movement states the only way to reduce further ecological destruction is to move away from the current capitalist system (Beer, 2022). However, due to a rise in greenwashing, new models such as Ecological Modernisation are unlikely to adequately address and halt the ecological crisis. Such is the case with carbon markets and credits, where carbon offsetting is bought and traded for by members in the Global North, with continued exploitation of natural areas and people in the Global South (Croeser, 2021). Of course, society is faced with continued population increase, demanding greater food production and shelter. However, we must move away from a system that continues to rely on fossil fuels and implement actual green infrastructure, which is separated from profit gain.

What further worsens societal capability to adequately address the ecological crisis, is lack of ownership, denial, and disconnect from nature. The human/nature divide is not a new concept, and environmental educators and philosophers have linked this to the environmental issues we continue to face (Barry, 2009). Beery *et al.* (2023) outlined the elements of this disconnect, which has been worsened through removal of nature interaction and subsequent loss of knowledge, increasingly sedentary lifestyles, tech-focused, and urbanisation, as well as nature phobias and altering perspectives on what is natural and what is not (to name a few). These are exacerbated even more subtly in the depiction of nature in pop culture, for example (Prévot-Julliard *et al.,* 2015). On top of this, within the SDGs, adequate education of environmental issues and sustainability is not outlined within the relevant educational SDG4 (Walid and Luetz, 2018). This would suggest that being environmentally literate and biodiversity aware is not viewed as a key aspect of future sustainability. This intensifies the problem of effectively communicating the ecological crisis. This communication is hampered by a futuristic narrative surrounding the consequences of the crisis, lack of reward for action, and removed personalisation of the issues (Moser, 2010).

7.1.2 Environmental Ethics and Resource Use

There is contention as to whether society should base nature's value on the services it provides for us, rather than as an entity that has the right to exist outside of our own requirements. What is more, the value of these ecosystem services is largely linked to whether they are profitable or not, whilst some services (such as cultural) are difficult to measure, and, as such, put value on. For example, whilst provisioning services (e.g., food production) are increasing, other services, such as regulating (air quality) are worsening (Lant *et al.,* 2008). Our overuse of environmental resources associated with key services results in a case of Hardin's "Tragedy of the Commons" – whereby our assumption that certain ecosystem services are provided for free, results in the exploitation of these services (Stein, 2022). The general assumption that we will have continuous access to these ecosystem services will result in their unregulated and unsustainable use. However, the best way to express the gravity of the situation, and drive action towards the conservation of the planet, may be to focus the narrative on what the loss of these services means for humanity.

For action to occur, companies need to see the biodiversity crisis as a "business crisis", governments need to see it as a "political crisis", individuals need to see it as a "livelihood crisis", because it is all these things and an insurmountable amount more. As governments and businesses are driven by profit, it needs to become economically viable to operate pro-environmentally. Research is showing that individuals do consider a company's environmental stance and actions (Albrecht *et al.,* 2023). However, if individuals believe they have no power, they are unlikely to engage with pro-environmental behaviours, or action against governments and organisations. Yet, evidence suggests that governments and organisations do respond to public demand (Schaffer *et al.,* 2022; Zhou *et al.,* 2021). Therefore, it is undeniable that this must be a collaborative effort, and no one sector will be able to address the challenges alone.

Potentially, once the urgency has been grasped, we can design a new relationship with nature, eliminating the human-nature divide, and forming a holistic relationship, one removed from egocentricity. This follows Leopold's (1949) concept of "Land Ethic", where community is created between both humanity and the environment. Considerate interaction and use of natural resources will create a positive feedback loop, whereby the land can continue to support humanity (Leopold, 2004). This view is further supported by the idea of Deep Ecology, and that harmonious living may only occur through a shift in society's perspective on the natural world, and uptake of ecocentrism (Naess, 1973; Taylor *et al.,* 2020).

7.1.3 A Call for Implementing Change

We need to improve our techniques in mitigating and reversing the trends we are seeing across nature and the climate. Environmental monitoring is one of the first steps towards this, by providing critical baseline data, identifying priorities for targeted aims, assessing conservation success, and addressing consequent shortcomings. However, these efforts are driven by scientists and activists, not by governments, organisations, and the wider public. Without a collaborative effort and foresighted mindset, we will not be successful in reversing the damage that we knowingly continue to cause. Scientific research will continue to evidence our impacts, suggest solutions, and call for action, but this information needs transforming into daily efforts and governmental initiatives, and this must be included in all future scientific research objectives.

7.2 Development of the Thesis and Summary of Results

Through exploring the gaps in environmental monitoring, the literature highlighted; 1) lack of inclusion of certain habitats (global grasslands) and species (invertebrates), 2) biases in monitoring distribution (focuses on the Global North), and 3) deficiencies in science-into-action from the separation between researchers, policy makers, and the public. Methods for improved biodiversity monitoring were explored to address these gaps, specifically looking at citizen science (CS) and remote sensing (RS) (common and/or advancing tools in environmental monitoring) to widen spatial reach and address the highlighted gaps by working under an open science (OS) framework. These gaps allowed specific targets for environmental monitoring to be identified within the UK, addressing priority habitats and species whilst providing a framework for a methodology that can be applied more globally. It was a key aim of the thesis to enlist stakeholders and members of the public with the research to widen its impact for science, policy, and society. Collaboration with Butterfly Conservation was sought, as their conservation objectives are aligned to some of the highlighted biodiversity monitoring gaps. Initial discussions with Butterfly Conservation further defined the habitat (species-rich grasslands) and invertebrate (Northern Brown Argus Butterfly) targets, which aligned with the gaps identified from the literature review.

As a result, the thesis aims were then set out to; 1) determine the role of OS approaches in the conservation of biodiversity, and 2) combine two approaches in biodiversity monitoring, RS and CS, to create an open and interdisciplinary habitat monitoring tool. This was with the hopes of improving "bioliteracy", awareness of biodiversity issues in the UK, and conservation agency through the implementation of OS, help map a priority habitat for vulnerable species within the UK, and create a tool framework that could be used globally for biodiversity monitoring.

To be able to create this monitoring tool, a case-study focus was needed for the initial application of the research. Species-rich grasslands and *A. ataxerxes* were chosen as the thesis focus, as an underrepresented habitat in monitoring initiatives, with many associated vulnerable species, thus addressing RQ1. The thesis emphasised public engagement with biodiversity conservation, and, as such, it was paramount to work under an OS framework. It was, therefore, important to implement OS practices in the design of the tool, especially with CS being both an enabler and facilitator of OS. However, it was unclear how current biodiversity monitoring CS surveys adhered to these OS practices and could be embedded into the design of the surveys. As such, a systematic review was conducted to investigate if current biodiversity monitoring CS surveys adhere to OS practices, for greater scientific and societal impact. This research addressed RQ2 "To what extent do citizen science studies of biodiversity demonstrate the principles of open science?", guiding the creation of my own survey and future research. This exploration was the first of its kind to analyse a broader range of OS practices (not just the openness of data) in biodiversity monitoring CS surveys, especially when considering the use of data management plans and preregistrations. The research found that biodiversity monitoring CS surveys did not adhere well to OS practices, and the increase in the use of these surveys did not promote openness within the tool. The review highlights the associated issues with current surveys and provides guidance for new surveys to follow, to further the impact of these programmes for improved biodiversity monitoring attempts. Following on from this research, the recommendations listed were used in my own CS survey design and analysis of its openness (Table 7-1).

Table 7-1. Identified open science practices from chapter 2 and how these were implemented in the citizen science survey Ecosystem Explorers' research.

Identified Recommendations	Implementation in the 'Ecosystem Explorers' Survey
Provide an open data management plan	A data management plan was written and will be openly available with the publication of this thesis on The University of Glasgow`s Enlighten repository (http://eprints.gla.ac.uk/).
Create an open preregistration	A preregistration was created and published on the Open Science Framework (Suter *et al.*, 2023).
Use open softwares and data	Open Sentinel-2 satellite imagery was accessed via Google Earth Engine. R was used to extract reflectance values and create a habitat prediction model (both of which were used to create the outputs for the survey).
Ensure subsequent data and code is open	The data and code associated with this research will be published as supplementary materials alongside an open access journal publication.
Results must be open and accessible	The data collected from participants is found on the project platform (https://www.citsci.org/projects/ecosystem-explorers). Ethical requirements meant that only members of the project can access the data, however, anyone can sign up to become a member and view the data. Results of the CS survey will be published in an open access journal.

It was important to utilise advancements in RS to aid in the mapping and monitoring of the identified conservation targets of the research. This required detailed environmental and RS data collection of known SRG characteristics. Secondary data (on habitat coverage, climate, and species distributions) were explored to locate potential sites of SRGs to collect environmental data that would characterise

SRG classes. This data exploration and spatial analysis guided the field surveying methodology to locate survey sites, leading to the creation of a RS model that would predict areas of SRGs across Scotland. Previous classification schemes were synthesised and adapted to create my own classification schema, further informed by the environmental and RS data that were collected at the survey sites. This research addressed the thesis and case-study specific RQ3, "Can a habitat classification model be created to predict species-rich grasslands in Scotland and locate habitat for vulnerable species?", surveying known species-rich grassland sites in Scotland to collect environmental and RS data to classify these habitats and create a SRG habitat prediction model to be applied to satellite imagery nationally. As a result, five SRG classes were decided on due to differing species composition, which was reflected in the spectral profiles. A SRG habitat classification model was created with 98.6% accuracy on test data. This model was applied to wider satellite imagery across Scotland to predict possible locations of SRGs. The accuracy of extrapolating the model to a greater spatial reach across Scotland was assessed in chapter 6.

Due to the scale, structural, and spectral similarities of the various SRG classes, the possibility of other RS techniques were researched. Common applications of retrieving and predicting grassland traits were investigated specifically for these SRGs, to consider if these would further help habitat classification models in their predictions of these habitats, through their differentiation. Prediction estimates of species diversity (richness) and community traits (including structural and biochemical) were investigated through common modelling techniques using spectral diversity, surface reflectance values, and vegetation indices. This investigation addressed the thesis and case-study specific RQ4 "Is currently available open-source remote sensing data able to accurately monitor species-rich grasslands and their vulnerable species", by exploring RS applications, including the relationship between species diversity and spectral diversity, and the retrieval success of grassland community traits in species-rich grasslands across Scotland, to see if these could help future mapping of the habitats. The results found that RS components were largely poor predictors of the environmental features in SRG environments, both within and across sites. High variation within and between classes causes too many confounding influences that are difficult to account for. Recommendations were suggested for RS in these environments; however, this information may be better used in monitoring of fully assessed sites, rather than assisting future mapping attempts. It was not possible to use these applications to aid the habitat mapping of this thesis, but it is with hopes that this research provided new knowledge to the field of RS, especially in the context of grassland conservation.

Utilising the survey design initially designed with Butterfly Conservation and the outputs of the RS habitat mapping model applied to satellite imagery covering the whole of Scotland, I devised a linked CS survey, Ecosystem Explorers. Continuous discussions with Butterfly Conservation throughout the research process enabled me to refine my earlier methods into the CS survey. Through the combination of CS and RS, it was determined that, for the purposes of this thesis, the public would be engaged to help confirm the outputs of the habitat classification model predictions. The survey was implemented over two summer seasons and data was supplied either from confirmed areas of SRGs within the last two years, or from exploring predicted areas, both of which were compared to predicted classifications. This part of the research addressed the thesis and case-study specific RQ5 “Can citizen science data validate the outputs of remote sensing models to identify species-rich grasslands for vulnerable species protection?”. This involved the alignment of CS and RS data in one open access tool to identify areas of SRGs for priority species protection, using the CS data to validate the outputs of RS models, and assess the involvement of citizens in generating information on species-rich grassland mapping. OS was able to be employed largely throughout the research process and combine RS outputs in a CS survey. The resulting project is predicted to be highly open, and showcases how RS and CS can be combined in an interdisciplinary approach. Although RS data could be utilised in a CS survey for habitat mapping, the habitat classification model did not appear to successfully predict areas of SRGs on a large scale, due to poor alignment with the citizen ground-truthed data. It was found that participants’ plant identification experience and confidence did not impact the results of the CS survey and are valuable contributors to biodiversity conservation efforts. This was also not influenced by who provided the data (e.g., novice citizen scientists or veteran surveyors).

This research highlighted the need for increased open-source RS data to provide a global biodiversity observation network across all contexts. It demonstrates that working in an OS framework is not always possible for specific habitats and this may hamper progress in biodiversity conservation. However, the tool was able to successfully engage with stakeholders through continuous collaboration with Butterfly Conservation and brought in further partnerships with other NGOs (such as Plantlife, The Conservation Volunteers, and the Botanical Society of Britain and Ireland) during the implementation of the CS survey. Although many stakeholder groups were able to be enlisted in the survey, there is still large room for improvement in biodiversity conservation engagement, specifically surrounding ecosystem-based research. This was emphasised by the individual participant numbers. Nevertheless, from interactions in the field with participants, there is enthusiasm that can be harnessed, and this is where research should focus.

Overall, the research in this thesis was able to provide a methodology that combines CS and RS in a novel technique for habitat mapping (Figure 7-1). However, it identifies the downfalls of limited open access data. This tool has wider biodiversity applications, and it is likely that the poor extrapolated accuracy of a Sentinel-2 based SRG habitat prediction model would not deter future projects or researchers from investigating similar methodologies in other contexts, specifically for biodiversity conservation.

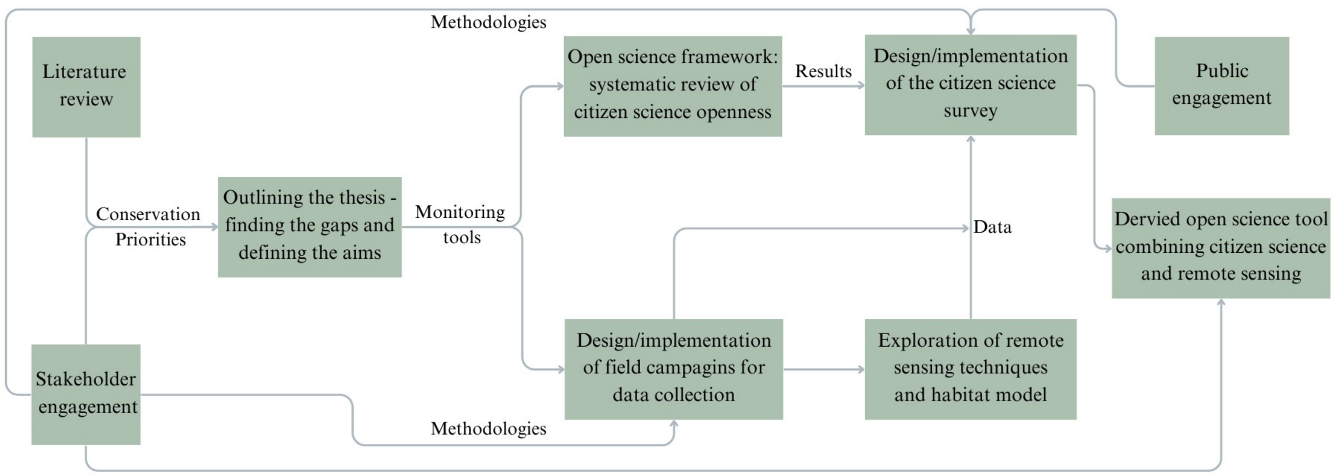


Figure 7-1. Process of designing the thesis elements and chapters.

7.3 Research Limitations

To work in an OS framework, S2 satellite data was chosen as a freely accessible data source for satellite imagery. Although it has enhanced benefits over other multispectral satellites, the spatial resolution of 10 m is too large for accurate predictions of Scottish SRGs. Due to the nature of the habitat, a much finer spatial resolution would be needed to capture the intra-class communities and avoid other/edge habitat interference. However, this could not be open source. Planetscope data (which has a higher spatial resolution of 3 m compared to Sentinel-2's 10 m), was acquired and a model could be created and compared to assess the extrapolated accuracy of predictions across Scotland. However, time restraints would not allow for this but would be an area for future research. This may be an issue related to SRGs. It would be relevant to investigate whether the tool could work in an OS approach for mapping other priority habitats, further discussed below.

Due to the available data online and time constraints for field work, the model was trained with 10 habitat classes (five SRGs and five background). In actuality, there are many more habitats defined across Scotland, and in the UK, than this; including, but not limited to, wetlands and heathlands, which could not be identified from aerial imagery like some of the other background classes (such as water and woodland). This means that there are likely overestimations of SRG predictions where heathlands may be classed as acid grassland, or wetlands may be classed as marshy grasslands, as each pixel must be classified in the image. The model would benefit from further training data on other habitat classes.

The CS survey is also limited to how many participants provided data to the project. Although attempts were made to increase involvement through a variety of means over the two summers, uptake was surprisingly low throughout. The number of participants also limited the spatial reach of the project; the habitat classification model predicted SRGs on a far greater spatial scale than the number of participants were able to cover in their ground-truthing. This low participation also meant that very little data on *A. artaxerxes* habitat was able to be collected. *H. nummularium* and subsequent *A. artaxerxes* egg surveys could not be conducted to provide further information to Butterfly Conservation. The project would need to run for much longer than the two summers that it was able to, to capture this information. These limitations could have been impacted by Scotland's small population size (5.4 million as of 2022) (Scotland Census, 2022). Potentially, if the tool had been instigated elsewhere (for example, to the whole of the UK), the CS survey may have gathered greater attention. This small population may also mean chance encounters between people and nature (for example, butterflies) is less likely, and the interest in surveys such as these may not exist widely throughout the country.

7.4 Wider Implications

This thesis has aimed to create an open, interdisciplinary tool for biodiversity monitoring in a time of ecological crisis. The combination of RS and CS has been called upon in recent years but has not yet achieved its apparent potential, especially in a biodiversity conservation perspective.

7.4.1 Addressing Species-Rich Grassland Protection in the UK

This novel tool was used in the context of SRGs, an under-mapped and -monitored habitat in Scotland. These habitats need identification and protection to continue to support pollinators and store carbon. The tool has been instigated across Scotland but is applicable to map SRGs across the UK through the input of higher spatial resolution RS data and model refinement: with the addition of further excluded habitat classes (e.g., peatlands, shrub, moorlands). Any further refined outputs and results from the research could help inform other government projects that include the mapping of SRGs, such as the Habitat Map of Scotland (Hester and Scholtens, 2019), the Living England Habitat Map (Natural England, 2022), and Living Wales (Planque *et al.,* 2022). Furthermore, if the RS model can be improved for SRGs, then the tool can go beyond mapping SRGs and enable the monitoring of short term (interannual-decadal) change at smaller spatial scales too. It was evident in the research conducted in chapter 4 that the success of trait retrieval varied per site each year (see Appendix D-2). This suggests that something has changed on site from one year to the next (weather/management/disease) and that it may be worth exploring.

The loss of SRGs over the last 60 years has led to conservation efforts to prevent and reverse their decline. For example, through their inclusion in protected areas and in agri-environmental schemes in the UK, where subsidies are provided to encourage farmers to manage their land in a biodiversity friendly manner. This enables specific grazing regimes (usually with cattle) to be employed on the habitat, which is vital to keep threats, such as bracken and scrub encroachment minimal but ensures no over-grazing, allowing seed set and flowering of essential flora (Hall, 2010). As much of the habitat is lost, protection of the remaining area, as well as restoration and creation, are essential. This has resulted in calls from environmental organisations of increased legal protection, as well as the restoration or creation of 12,000 ha of SRGs in the UK by 2043 (Plantlife, 2018). Such restoration efforts can be highly effective; where SRGs have previously been restored floral and insect diversity has improved comparative to ancient SRG sites (Auestad *et al.*, 2016; Dicks, 2002; Forup and Memmott, 2005). Similarly, inclusion or creation of SRGs in agri-environmental schemes has the potential to benefit insect numbers greater than that of field margins, hedgerows, or even some ancient SRGs, depending on the time of year (Lye *et al.*, 2009).

It is apparent that SRG protection and restoration can benefit these habitats and aid declining associated invertebrate populations, but this is not always the case. Analysis of SRG sites in agri-environmental schemes between 2006 to 2014 showed that only 23% of sites had improved in their condition, with 11% declining (Peel, 2017). Results have varied in their successes due to soil fertility, location, restoration and management methods, and landowner experience and incentive (Warwickshire County Council, 2018). It is suggested that improvements can be made with site specific management and increased contact between experts and landowners (Peel, 2017). On an individual scale, it is evident of variation in site and study outcome, however, what is apparent is that these methods are not being done on a large or fast enough scale to be widely beneficial to the UK or even European landscape. Monitoring is essential to assess the effectiveness of these methods and the current extent and condition of SRGs. Therefore, under-monitoring of SRGs and associated species must be addressed to preserve these essential ecosystem components.

The methodologies associated with Ecosystem Explorers could be applied throughout Europe with the addition of extra grassland classes, such as those outlined in EUNIS, and further synthesis between classification schemes. The tool could assist in mapping other priority habitats (such as peatlands within the UK) both nationally and globally, as it has shown how the public can interact with RS data and how organisations are able to utilise the results to reach their own targets. It would be necessary to adapt the inputted RS data to the targeted habitat. As it was discussed in chapter 5, the spatial and spectral scale of the RS sensor will be dictated by the habitat features. For example, mapping peatland vegetation has included the use of hyperspatial, multispectral, and Synthetic Aperture Radar sensors (Bourgeau-Chavez *et al.*, 2017; Chimner *et al.*, 2019; Frick *et al.*, 2011). Gamon *et al.* (2020) specifically highlight that any global biodiversity monitoring (or in this case mapping) tool, must be “scale-aware”, and there is no one-size-fits all, further supporting a flexible approach to any use of RS data for this purpose.

7.4.2 Global Implementation and Targeting Excluded Groups

The research in this thesis has provided an OS framework to guide projects, not limited to CS, in complying with OS practices, increasing their transparency and replicability. It also demonstrates how to involve organisations, governmental workers, and the public in scientific research to utilise their specialised knowledge and engage their interest in important biodiversity concerns, which could hopefully be applied to other conservation targets. It was identified in chapter 1 that many CS programmes and biodiversity monitoring programmes are not found in the Global South. When addressing the gaps highlighted in the literature review, the research in this thesis was not able to target all monitoring gaps but it was with hopes that the derived tool could be instigated in the Global South, for example.

It is known that CS projects have struggled to engage all demographics (biases are seen in ethnicities, gender, and socio-economic class) in the past; so much so that even the term "CS" is under contention (Liebenberg *et al.,* 2021; Pateman *et al.,* 2021*a*). The lower costs associated with CS for biodiversity monitoring are important to take advantage of in developing countries and enlist the help from traditionally excluded demographic groups. The use of technology may allow people to participate with reduced mobility or by removing spatial barriers, and to contribute from home via computers or in their local region via their phones. RS data from certain satellites (such as Sentinel-2) is freely available, meaning that this data can be acquired in areas where more costly data cannot be applied. The limitations here would be access to computers to make use of this data, known as "digital exclusion" (Pateman *et al.,* 2021*b*). In areas where this is not possible, participatory research may be vital to ensure equality and inclusion in biodiversity monitoring (Davis *et al.,* 2020).

However, there are still issues with global collaboration; Heberling *et al*. (2020) found that 8% of biodiversity studies in the Global South did not have proportional representation of locals at author level. Although this relationship was found to be improving, it is with hopes that OS tools, such as Ecosystem Explorers can address this with greater involvement of local populations. Gaining perspectives and knowledge of local people in the Global South through interviews and focus groups will integrate information on data that needs to be collected (Danielsen *et al.,* 2018), which can then be realised through RS targeting. Where language or literacy barriers exist, the use of pictures or translators to display information in another form may be possible as has occurred in UCL's The Extreme CS research group, whilst hard copy data collection methods are still viable with pen and paper (Paleco *et al.,* 2021; Stevens *et al.,* 2014).

The possibility when pursuing inequality in biodiversity monitoring is to specifically direct these methods at excluded groups (Haklay, 2015). It was with hopes that the methodology for participating in Ecosystem Explorers was as accessible as possible, requiring little to no equipment, whilst survey days were offered around Scotland to reach local groups who could not travel far. The survey days also involved training aspects of plant identification and surveying methods to encourage people from all backgrounds no matter their employment or educational status. This allowed members of the public including children and those in routine manual and service jobs to access training that is usually costly (pers.obvs.). It is observed in chapter 5 that all demographics were not reached, and, in the future, it would be a specific objective to address this issue through the discussed methods above.

7.4.3 As an Educational Tool

If there is increased pressure from the public that has been bred from knowledge on biodiversity issues, there may be more of an inclination to meet the biodiversity targets that are constantly set and rarely achieved (Harlin *et al.,* 2018). Increasing biodiversity awareness needs to be implemented across all stages of life, especially in younger children, as this will hopefully breed a respect and interest that will continue over the years. Kelemen-Finan *et al.* (2018) found that primary aged children reached their learning targets more readily than older school students where CS was part of their education. This further supports the need to increase scientific literacy at a young age. As discussed in chapter 1, CS projects can increase knowledge, such as species identification, but this rarely translates into behavioural changes outside of the project. Evidence suggests that practical and research project-based education is important to create this behavioural change, so developing a CS project which provides more than just the possibility of data collection could be a way to manage this (Saunders *et al.,* 2018). For example, providing agency within projects should allow participants to take home what they have learnt and include these actions in their everyday life (Kılınç, 2010; Yli-Panula *et al.*, 2018).

It has been noted that one of the most difficult challenges of integrating RS in CS and biodiversity are the knowledge gaps. Open access projects, such as Ecosystem Explorers, may facilitate the integration and understanding across the disciplines through public participation and interaction with RS outputs in the survey, as well as access to open data and code for replicability in schools, for example. Programmes, such as SatSchool (2022) now allow accessible training in RS, and this should further support incorporation of RS data into ecology research. Issues around the training would largely be linked to lack of uniformity across RS approaches but free tools, such as GIS.lab, allows the comparison of approaches across software packages by running simultaneously (Rocchini *et al.,* 2017).

It was an initial aim of this thesis to integrate the mapping tool, or Ecosystem Explorers, into schools to address curriculum targets in Biology, Geography, and IT. However, while meetings were held with the local authority and associated STEM education groups, time limitations, and the COVID-19 pandemic, meant this was not developed to the planned extent. By implementing the tool into curriculums there is wide potential to 1) increase engagement with active scientific research, 2) expand knowledge on ecological RS, mapping and geospatial analysis, and species identification, and 3) potentially enhance external environmental interaction and conservation proactivity. For example, the CS survey ClimateWatch was used in a university curriculum, which resulted in greater environmental engagement and improved scientific precision (Mitchell *et al.,* 2017). Implementation in schools has benefits, not only for biodiversity conservation, but on an educational and personal level too. Engaging with outdoor, nature-based, or CS activities has been shown to improve learning and development through heightened cognition, physical health through increased activity and movement, and mental wellbeing from elevated mood of pupils (Booth *et al.,* 2020; Harvey *et al.,* 2020; The Newport Biodiversity in Schools Project, 2011). This implementation will, therefore, be mutually beneficial for educational and environmental progress (Figure 7-2).

Figure 7-2. Educational and Environmental goals that can be reached through the implementation of biodiversity citizen science and remote sensing programmes in schools.

The tool could be implemented in the curriculum through 1) promotion to schools or the educational estate to gain interest, 2) initial discussions with participating schools, 3) the development of activities across year groups per targeted subject, with teacher training events for success, 4) implementation of activities, 5) feedback and evaluation of activities, and 6) assessment of students' increased literacy of the subject and subsequent behavioural change (Figure 7-3).

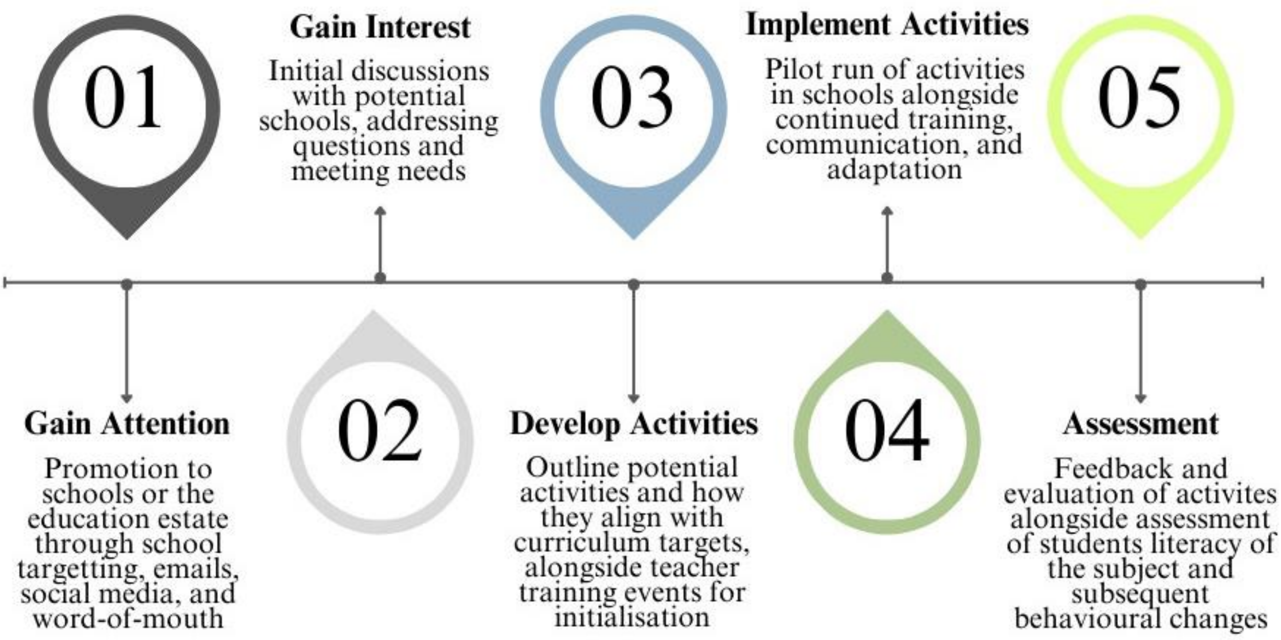


Figure 7-3. The process of implementing projects, such as Ecosystem Explorers, into schools and the curriculum.

The initial outlines of project activities and how they achieve curriculum targets for schools can be found in Appendix F-1. Future research of the tool's implementation in curriculums is vital for addressing the concerns of the loss of, not only science literacy, but specifically plant literacy that is being seen (Shah and Martinez, 2016; Stroud *et al.,* 2022).

7.5 Future Research

It is widely accepted that RS of SRGs is challenging, however, with continuing advancements in RS technology, this is predicted to improve in the future. Specifically, higher spatial resolution data (such as the data from Planetscope or through exploration of WorldView-2 data) is more likely capable of capturing data at species level and could be used in the classification model to improve the accuracy of the habitat predictions. In addition, a hierarchical scheme could be used to predict higher classifications of SRG grasslands, for example, NVC communities within the broader classes defined in this thesis. The Micasense data may more effectively pick up differences between these communities and it would be useful to see how RS data across spatial and spectral scales predict multi-level grassland classifications.

The issue with using 10 m spatial resolution satellite imagery highly affected the potential to target Butterfly Conservation's priority species' (*A. artaxerxes*) food plant, *H. nummularium*. RS data has previously been used to create spectral profiles of specific species for agriculture, invasive, and wild plant species (Khderry and Yones, 2021; Iqbal *et al.,* 2021). Future research could create a spectral profile for *H. nummularium,* both *in situ* and individually grown (which would provide comparisons) and investigate the retrieval of this information by upscaling it to predicted areas of SRGs. Research could explore if RS methods can more accurately identify the species in the broader habitats with defined spectral signatures of food plants for priority species protection.

The method is also one example of how RS data can be used in CS surveys, but it is not limited to only this input (of ground-truthing). To target one of the limitations outlined, data on other background habitats could be provided for by members of the public, rather than relying on further field work or secondary data. It would be unique to investigate the accuracy of reference/train data supplied for by members of the public. Furthermore, research could further explore how to continue the increase of scientific democratisation, and if this process has had any consequence on plant literacy and conservation agency of its participants. Similarly, the development of OS educational resources and a common framework for practicing OS could be investigated. The success of this could be evaluated through exploring the uptake of practices and how this has affected scientific practice and research impact. This is crucial for continued progress in this area, and the implementation of OS practices in research and CS in early education should be pursued. Not only this, but surveying of the participants understanding of RS data, and the interaction with it, can assess how successful the integration of disciplines was for increased public understanding and for further enhancing biodiversity conservation.

7.6 Summary and Concluding Remarks

The research in this thesis aimed to target highlighted gaps in biodiversity monitoring by using tools that have applications in ecology observations (CS) and land cover assessment (RS), under an OS framework. The research was directed at species and habitat biases by focusing on SRGs and associated invertebrate species. The research also contributed to the limited applications of combining CS and RS in a cross-disciplinary approach, to bridge knowledge gaps and increase data collection capacity for biodiversity conservation. OS practices were in place throughout, such as the use of open software and open data, whilst any outputs from this thesis were made, or will be made when applicable, open as well (through open access publications, a preregistration, and associated openly published data and code).

The research emphasises that the openness of research within biodiversity monitoring is not at the appropriate level for the chosen methodologies. It also stresses how limited open access data may reduce the success of RS applications in the context of SRGs, and further improvements in the technologies are needed for these habitats. It must be noted that there is potential for RS to enhance biodiversity monitoring and habitat mapping, however, models are just that – models. The research highlights that RS data, or any technological assistance, should not be in replacement of having *in situ* verification and should be more of an accompaniment. It also provides evidence that citizen scientists can provide ground-truthing data of comparable value to that of professional surveyors. The continuation of human involvement in biodiversity monitoring not only assures high quality data, but also keeps humans interacting with nature – to let them increase their bio literacy, see the changes (both positive and negative), and increase agency for biodiversity conservation; both for the planet’s benefit and our own.

References

Albrecht, S.L., Donnelly, T., Frenkiel, M., Rajic, S.K., Kavadas, V. and Leiter, M.P., 2023. Pro-environmental employee engagement: the influence of pro-environmental psychological capital, pro-environmental job resources, and perceived corporate environmental responsibility. *Frontiers in Sustainability*, *4*, p.1117892.

Auestad, I., Rydgren, K. and Austad, I., 2016. Near-natural methods promote restoration of species-rich grassland vegetation—revisiting a road verge trial after 9 years. *Restoration Ecology*, *24*(3), pp.381-389.

Barry, C., 2009. The environment/society disconnect: An overview of a concept tetrad of environment. *The Journal of Environmental Education*, *41*(2), pp.116-132.

Beer, C.T., 2022. "Systems change not climate change": support for a radical shift away from capitalism at mainstream US climate change protest events. *The Sociological Quarterly*, *63*(1), pp.175-198.

Beery, T., Stahl Olafsson, A., Gentin, S., Maurer, M., Stålhammar, S., Albert, C., Bieling, C., Buijs, A., Fagerholm, N., Garcia-Martin, M. and Plieninger, T., 2023. Disconnection from nature: Expanding our understanding of human-nature relations. *People and Nature*, *5*(2), pp.470-488.

Booth, J.N., Chesham, R.A., Brooks, N.E., Gorely, T. and Moran, C.N., 2020. A citizen science study of short physical activity breaks at school: improvements in cognition and wellbeing with self-paced activity. *BMC Medicine*, *18*(1), pp.1-11.

Bourgeau-Chavez, L.L., Endres, S., Powell, R., Battaglia, M.J., Benscoter, B., Turetsky, M., Kasischke, E.S. and Banda, E., 2017. Mapping boreal peatland ecosystem types from multitemporal radar and optical satellite imagery. *Canadian Journal of Forest Research*, *47*(4), pp.545-559.

Chimner, R.A., Bourgeau-Chavez, L., Grelik, S., Hribljan, J.A., Clarke, A.M.P., Polk, M.H., Lilleskov, E.A. and Fuentealba, B., 2019. Mapping mountain peatlands and wet meadows using multi-date, multi-sensor remote sensing in the Cordillera Blanca, Peru. *Wetlands*, *39*, pp.1057-1067.

Croeser, E., 2021. What to expect from COP26: climate action, climate justice, or green washing? *Institute for Interdisciplinary Research into the Anthropocene*. *[Online]*. Accessed 4th September 2023. Available at: https://iiraorg.com/2021/11/05/what-to-expect-from-cop26-climate-action-climate-justice-or-greenwashing/

Danielsen, F., Burgess, N.D., Coronado, I., Enghoff, M., Holt, S., Jensen, P.M., Poulsen, M.K. and Rueda, R.M., 2018. The value of indigenous and local knowledge as citizen science. UCL Press.

Davis, L.F., Ramírez-Andreotta, M.D. and Buxner, S.R., 2020. Engaging diverse citizen scientists for environmental health: Recommendations from participants and promotoras. *Citizen Science: Theory and Practice*, *5*(1).

Dicks, L.V., 2002. *The structure and functioning of flower-visiting insect communities on hay meadows* (Doctoral dissertation, University of Cambridge).

Forup, M.L. and Memmott, J., 2005. The restoration of plant-pollinator interactions in hay meadows. *Restoration Ecology*, *13*(2), pp.265-274.

Frick, A., Steffenhagen, P., Zerbe, S., Timmermann, T. and Schulz, K., 2011. Monitoring of the vegetation composition in rewetted peatland with iterative decision tree classification of satellite imagery. *Photogrammetrie-Fernerkundung-Geoinformation*, pp.109-122.

Gamon, J.A., Wang, R., Gholizadeh, H., Zutta, B., Townsend, P.A. and Cavender-Bares, J., 2020. Consideration of scale in remote sensing of biodiversity. *Remote Sensing of Plant Biodiversity*, pp.425-447.

Garthwaite, J., 2018. When did humans start influencing biodiversity? Earlier than we thought. *Stanford Earth Matters [Online]*. Accessed 27th July 2023. Available at: https://earth.stanford.edu/news/when-did-humans-start-influencing-biodiversity-earlier-we-thought

Haklay, M., 2015. *Citizen science and policy: A European perspective*. Washington, DC: Woodrow Wilson International Center for Scholars.

Hall, C., 2010. Management of species rich grasslands. Technical Note TN629. Available at: https://www.sruc.ac.uk/media/l1omdteo/tn629-rich-grasslands.pdf

Harlin, J., Kloetzer, L., Patton, D. and Leonhard, C., 2018. Turning students into citizen scientists. UCL Press.

Harvey, D.J., Montgomery, L.N., Harvey, H., Hall, F., Gange, A.C. and Watling, D., 2020. Psychological benefits of a biodiversity-focussed outdoor learning program for primary school children. *Journal of Environmental Psychology*, *67*, p.101381.

Heberling, J.M., Miller, J.T., Noesgaard, D., Weingart, S.B. and Schigel, D., 2021. Data integration enables global biodiversity synthesis. *Proceedings of the National Academy of Sciences*, *118*(6), p.e2018093118.

Hester, A.J. and Scholtens, R., 2019. *Analysing habitat patch size and connectivity at a national level for Scotland-data assessment*. James Hutton Institute, *14pp.*

Iqbal, I.M., Balzter, H. and Shabbir, A., 2021. Identifying the Spectral Signatures of Invasive and Native Plant Species in Two Protected Areas of Pakistan through Field Spectroscopy. *Remote Sensing*, *13*(19), p.4009.

Kelemen-Finan, J., Scheuch, M. and Winter, S., 2018. Contributions from citizen science to science education: an examination of a biodiversity citizen science project with schools in Central Europe. *International Journal of Science Education*, *40*(17), pp.2078-2098.

Khdery, G. and Yones, M., 2021. Innovative spectral library for identification common wild plants using hyperspectral technology in Northwestern Coast, Egypt. *The Egyptian Journal of Remote Sensing and Space Science*, *24*(3), pp.485-491.

Kılınç, A., 2010. Can project-based learning close the gap? Turkish student teachers and proenvironmental behaviours. *International Journal of Environmental and Science Education, 5*(4), pp. 495-509.

Lant, C.L., Ruhl, J.B. and Kraft, S.E., 2008. The tragedy of ecosystem services. *BioScience*, *58*(10), pp.969-974.

Leadley, P., Gonzalez, A., Obura, D., Krug, C.B., Londoño-Murcia, M.C., Millette, K.L., Radulovici, A., Rankovic, A., Shannon, L.J., Archer, E. and Armah, F.A., 2022. Achieving global biodiversity goals by 2050 requires urgent and integrated actions. *One Earth*, *5*(6), pp.597-603.

Leopold, A.C., 1949. *A Sand County Almanac and Sketches Here and There*. Oxford University Press, New York. Leopold, Luna B., ed., 1953.

Leopold, A.C., 2004. Living with the land ethic. *BioScience*, *54*(2), pp.149-154.

Liebenberg, L., Lombard, M., Shermer, M., Xhukwe, U., Biesele, M., Carruthers, P., Kxao, O., Hansson, S.O., Langwane, H.K., Elbroch, L.M. and Keeping, D., 2021. Tracking science: An alternative for those excluded by citizen science. *Citizen Science: Theory and Practice*, *6*(1).

Lye, G., Park, K., Osborne, J., Holland, J. and Goulson, D., 2009. Assessing the value of Rural Stewardship schemes for providing foraging resources and nesting habitat for bumblebee queens (Hymenoptera: Apidae). *Biological Conservation*, *142*(10), pp.2023-2032.

Mitchell, N., Triska, M., Liberatore, A., Ashcroft, L., Weatherill, R. and Longnecker, N., 2017. Benefits and challenges of incorporating citizen science into university education. *PLoS One*, *12*(11), p.e0186285.

Moser, S.C., 2010. Communicating climate change: history, challenges, process and future directions. *Wiley Interdisciplinary Reviews: Climate Change*, *1*(1), pp.31-53.

Naess, A., 1973. The Shallow and the Deep, Long Range Ecology Movements: A Summary. *Inquiry*, 16, p. 95–100.

Natural England., 2022. Living England Habitat Map (Phase 4). Accessed 20th September 2023. Available at: https://www.data.gov.uk/dataset/e207e1b3-72e2-4b6a-8aec-0c7b8bb9998c/living-england-habitat-map-phase-4#licence-info

Paleco, C., García Peter, S., Salas Seoane, N., Kaufmann, J. and Argyri, P., 2021. Inclusiveness and diversity in citizen science. *The Science of Citizen Science*, *261*.

Pateman, R.M., Dyke, A. and West, S.E., 2021(a). The diversity of participants in environmental citizen science. *Citizen Science: Theory and Practice*.

Pateman, R.M, Tuhkanen, H. and Cinderby, S., 2021(b). Citizen Science and the Sustainable Development Goals in Low and Middle Income Country Cities. *Sustainability*, *13*(17), p.9534.

Peel, S., 2017. Past and potential future roles of agrienvironment schemes. Accessed 8th March 2021. Available at: http://www.magnificentmeadows.org.uk/assets/uploads/11_Past_and_Potential_Roles_of_Agri-environment_Schemes_.pdf

Planque, C., Punalekar, S., Lucas, R., Chognard, S., Owers, C.J., Clewley, D., Bunting, P., Sykes, H. and Horton, C., 2020, September. Living Wales: automatic and routine environmental monitoring using multi-source Earth observation data. In *Earth Resources and Environmental Remote Sensing/GIS Applications XI* (Vol. 11534, pp. 41-49). SPIE.

Plantlife., 2018. Grasslands action plan. *[Online]*. Accessed 7th December 2020. Available at: https://www.plantlife.org.uk/application/files/2315/3087/2058/Grasslands_action_plan_-_Plantlife.pdf

Prévot-Julliard, A.C., Julliard, R. and Clayton, S., 2015. Historical evidence for nature disconnection in a 70-year time series of Disney animated films. *Public Understanding of Science*, *24*(6), pp.672-680.

Rathi, A., 2016. A 1912 news article ominously forecasted the catastrophic effects of fossil fuels on climate change. *Quartz [Online]*. Accessed 27th July 2023. Available at: https://qz.com/817354/scientists-have-been-forecasting-that-burning-fossil-fuels-will-cause-climate-change-as-early-as-1882

Rocchini, D., Petras, V., Petrasova, A., Horning, N., Furtkevicova, L., Neteler, M., Leutner, B. and Wegmann, M., 2017. Open data and open source for remote sensing training in ecology. *Ecological Informatics*, *40*, pp.57-61.

SatSchool., 2022. Available at: https://satschool-outreach.github.io/

Saunders, M.E., Roger, E., Geary, W.L., Meredith, F., Welbourne, D.J., Bako, A., Canavan, E., Herro, F., Herron, C., Hung, O. and Kunstler, M., 2018. Citizen science in schools: Engaging students in research on urban habitat for pollinators. *Austral Ecology*, *43*(6), pp.635-642.

Schaffer, L.M., Oehl, B. and Bernauer, T., 2022. Are policymakers responsive to public demand in climate politics?. *Journal of Public Policy*, *42*(1), pp.136-164.

Scotland Census, 2022. Accessed 2nd October 2023. Available at: https://www.scotlandscensus.gov.uk/2022-results/scotland-s-census-2022-rounded-population-estimates/

Shah, H.R. and Martinez, L.R., 2016. Current approaches in implementing citizen science in the classroom. *Journal of Microbiology & Biology Education*, *17*(1), pp.17-22.

Stein, G.M., 2022. Environmental Justice and the Tragedy of the Commons. *Calif. L. Rev. Online*, *13*, p.10.

Stevens, M., Vitos, M., Altenbuchner, J., Conquest, G., Lewis, J. and Haklay, M., 2014. Taking participatory citizen science to extremes. *IEEE Pervasive Computing*, *13*(2), pp.20-29.

Stroud, S., Fennell, M., Mitchley, J., Lydon, S., Peacock, J. and Bacon, K.L., 2022. The botanical education extinction and the fall of plant awareness. *Ecology and Evolution*, *12*(7), p.e9019.

Suter, S., Welden, N., & Barrett, B., 2023. A citizen science tool for monitoring species-rich grasslands and associated species. https://doi.org/10.17605/OSF.IO/WVQZS

Taylor, B., Chapron, G., Kopnina, H., Orlikowska, E., Gray, J. and Piccolo, J.J., 2020. The need for ecocentrism in biodiversity conservation. *Conservation Biology*, *34*(5), pp.1089-1096.

The Newport Biodiversity in Schools Project., 2011. Accessed 20th September 2023. Available at: https://www.newport.gov.uk/documents/Schools-and-Education/Biodiversity-in-Schools-Project.pdf

van Goethem, T. and van Zanden, J.L., 2021. Biodiversity trends in a historical perspective. in *How Was Life? Volume II: New Perspectives on Well-being and Global Inequality since 1820*, OECD Publishing, Paris, https://doi.org/10.1787/2c94883d-en.

Walid, M. and Luetz, J.M., 2018. From education for sustainable development to education for environmental sustainability: reconnecting the disconnected SDGs. *Handbook of Sustainability Science and Research*, pp.803-826.

Warwickshire County Council., 2018. Scientific evidence for habitat creation and restoration. *[Online]*. Accessed 8th March 2021. Available at: https://api.warwickshire.gov.uk/documents/WCCC-863-794

Weston, P., 2022. Humans v nature: our long and destructive journey to the age of extinction. *The Guardian [Online]*. Accessed 27th July 2023. Available at: https://www.theguardian.com/environment/2022/nov/25/cop15-humans-v-nature-our-long-and-destructive-journey-to-the-age-of-extinction-aoe

Yli-Panula, E., Jeronen, E., Lemmetty, P. and Pauna, A., 2018. Teaching methods in biology promoting biodiversity education. *Sustainability*, *10*(10), p.3812.

Zhou, J., Sawyer, L. and Safi, A., 2021. Institutional pressure and green product success: the role of green transformational leadership, green innovation, and green brand image. *Frontiers in Psychology*, *12*, p.704855.

Appendices

Appendix A

Appendix A-1. Citation for Northern Brown Argus occurrence data downloaded from NBN Atlas

Citation: NBN Atlas occurrence download at https://nbnatlas.org accessed on Wed Dec 02 15:47:24 UTC 2020
Records provided by Derbyshire Biological Records Centre, accessed through NBN Atlas website.
Records provided by Derbyshire Wildlife Trust, accessed through NBN Atlas website.
Records provided by ERIC NE Combined dataset to 2017, accessed through NBN Atlas website.
Records provided by Wildwatch North Pennines AONB project records for Cumbria, accessed through NBN Atlas website.
Records provided by Natural Resources Wales, accessed through NBN Atlas website.
Natural Resources Wales. (2020). Welsh Invertebrate Database (WID). Occurrence dataset accessed through the NBNAtlas
Records provided by North East Scotland Biological Records Centre, accessed through NBN Atlas website.
Natural England. (2020). Invertebrate Site Register - England (1738-2005). Occurrence dataset on the NBN Atlas
Casual records for Scottish Wildlife Trust reserves - Verified data (2020)
Records provided by Fife Nature Records Centre, accessed through NBN Atlas website.
Records provided by Argyll Biological Records Centre, accessed through NBN Atlas website.
Records provided by The Wildlife Information Centre, accessed through NBN Atlas website.
Records provided by Buglife, accessed through NBN Atlas website.
Records provided by Scottish Wildlife Trust, accessed through NBN Atlas website.
Argyll Biological Records Centre (2020). Argyll Biological Records Dataset
Records provided by Natural England, accessed through NBN Atlas website.
Contains UK Butterfly Monitoring Scheme (UKBMS) data Â© copyright and database right Butterfly Conservation, the Centre for Ecology & Hydrology, British Trust for Ornithology, and the Joint Nature Conservation Committee
Records provided by Caledonian Conservation, accessed through NBN Atlas website.
Records provided by Highland Biological Recording Group, accessed through NBN Atlas website.
Lothian Wildlife Information Centre surveys (Invertebrates - general), TWIC (2020)
Records provided by Environmental Records Information Centre North East, accessed through NBN Atlas website.
Records provided by UK Butterfly Monitoring Scheme, accessed through NBN Atlas website.
Records provided by National Trust, accessed through NBN Atlas website.
Records provided by National Trust for Scotland, accessed through NBN Atlas website.
Records provided by Butterfly Conservation, accessed through NBN Atlas website.
Records provided by SER Species-based Surveys, accessed through NBN Atlas website.
Records provided by National Trust Species Records, accessed through NBN Atlas website.
Records provided by Tullie House Museum Natural History Collections, accessed through NBN Atlas website.
Records provided by Cumbria Biodiversity Data Centre, accessed through NBN Atlas website.
Records provided by Fife Nature Records Centre combined dataset, accessed through NBN Atlas website.
National Trust for Scotland Species Records, NTS (2019)
Records provided by Lancashire Environment Record Network, accessed through NBN Atlas website.
Survey and monitoring records for Scottish Wildlife Trust reserves from reserve convenors and Trust volunteers - Verified data, SWT (2020)
Commissioned surveys and staff surveys and reports for Scottish Wildlife Trust reserves - Verified data, SWT (2020)

TWIC Biodiversity Field Trip Data (1995-present), TWIC (2019)
Highland Biological Recording Group (2020). HBRG Insects Dataset. Occurrence dataset accessed through the NBN Atlas
Records provided by Derbyshire Casual LEPIDOPTERA records - Casual records collated by Derby Museum., accessed through NBN Atlas website.
Butterfly distribution data from the Butterflies for the New Millennium recording scheme, courtesy of Butterfly Conservation and the Biological Records Centre.
Records provided by Staffordshire Wildlife Trust Nature Reserves Inventory, accessed through NBN Atlas website.
Records provided by NE Scotland butterfly and moth records 1800-2010, accessed through NBN Atlas website.
Records provided by Lepidoptera Records up to April 2010, accessed through NBN Atlas website.
Records provided by LERN Records, accessed through NBN Atlas website.
Invertebrate records from sites that are mainly across Scotland, Buglife (2020)
Natural England iRecord Surveys (2020)
Records provided by Staffordshire Ecological Record, accessed through NBN Atlas website.
Contains UK Butterfly Monitoring Scheme (UKBMS) data Â© copyright and database right Butterfly Conservation, the Centre for Ecology & Hydrology, British Trust for Ornithology, and the Joint Nature Conservation Committee.
Caledonian Conservation Ltd

Appendix A-2. Citation for Common Rock-rose occurrence data downloaded from NBN Atlas.

Citation: NBN Atlas occurrence download at https://nbnatlas.org accessed on Fri Dec 11 11:51:59 UTC 2020.
Records provided by National Trust for Scotland, accessed through NBN Atlas website.
Records provided by Botanical Society of Britain & Ireland, accessed through NBN Atlas website.
TWIC General Records (2015 - present), TWIC (2020)
Natalie Harmsworth's Records (2010-2019), TWIC (2019)
National Trust for Scotland Species Records, NTS (2019)
Records provided by Fife Nature Records Centre, accessed through NBN Atlas website.
Survey and monitoring records for Scottish Wildlife Trust reserves from reserve convenors and Trust volunteers - Verified data, SWT (2020)
Commissioned surveys and staff surveys and reports for Scottish Wildlife Trust reserves - Verified data, SWT (2020)
Records provided by St Andrews BioBlitz 2014, accessed through NBN Atlas website.
TWIC Biodiversity Field Trip Data (1995-present), TWIC (2019)
Records provided by Argyll Biological Records Centre, accessed through NBN Atlas website.
Records provided by The Wildlife Information Centre, accessed through NBN Atlas website.
Records provided by St Andrews BioBlitz 2016, accessed through NBN Atlas website.
Records provided by John Muir Trust, accessed through NBN Atlas website.
Records provided by Scottish Wildlife Trust, accessed through NBN Atlas website.
Argyll Biological Records Centre (2020). Argyll Biological Records Dataset
TWIC Site Surveys (2010 - present) (2020)
Records provided by NatureScot, accessed through NBN Atlas website.
SNH (2020). Grassland Surveys of Fife, 1972-1990
SNH (2020). Grassland Surveys in North-East Scotland, 1989
Botanical Society of Britain and Ireland. [2020.] Vascular plant records verified via iRecord.
Records provided by Scottish SNH-funded BSBI records, accessed through NBN Atlas website.
Records provided by Plants and Bryophytes recorded on Schiehallion 25-30 June 2000, accessed through NBN Atlas website.
Records provided by St Andrews BioBlitz 2015, accessed through NBN Atlas website.
Records provided by Other BSBI Scottish data up to 2015, accessed through NBN Atlas website.
Records provided by City of Edinburgh Natural Heritage Service - Historical Records, accessed through NBN Atlas website.

Appendix B – Chapter 2 Supplementary Materials

Appendix B-1 – R Code

````
##Load Packages

```{r}

library("tidyverse")

library("ggplot2")

library("lme4")

library("ordinal")

library("MASS")

```

##Load Data

```{r}

df <- read.csv(".csv")

```

##Is there a significant difference in openness scores across principles?

```{r}

kruskal.test(Score ~ Principle, data = df)

```

##Which principles significantly differ from each other?

```{r}

dunnTest(Score ~ Principle,

         data=df,

         method="bonferroni")

```

##Investigating change in adherence to CS over time

##Mutate score to factor

```{r}

df <- df %>%

  mutate(score = factor(score))

```

##Rescale variables

```{r}

df$year <- scale(df$year, center = TRUE, scale = TRUE)

```

##Ordinal Regression model for change in average openness in publications over time

```{r}

ordinalmodel <- clm(openness_scores ~ publication_year, data = df, link = "logit")
````
````

````
summary(ordinalmodel)

```

##Ordinal Regression model for change in openness scores per principle in publications over time

```{r}

ordinalmodel <- clm(openness_within_principle ~ publication_year, data = df, link = "logit")

summary(ordinalmodel)

```
````
````

Appendix B-2. The 153 studies identified after the methods sorting. The final 42 studies utilised for all analysis are highlighted in green.

Authors	Article Title	Publication Year
Becker, CD; Agreda, A; Astudillo, E; Costantino, M; Torres, P	Community-based monitoring of fog capture and biodiversity at Loma Alta, Ecuador enhance social capital and institutional cooperation	2005
Pieterson, E. Corrie; Addison, Lindsay M.; Agobian, Jorge N.; Brooks-Solveson, Brenda; Cassani, John; Everham, Edwin M., III	Five years of the Southwest Florida Frog Monitoring Network: Changes in frog communities as an indicator of landscape change	2006
Goffredo, Stefano; Pensa, Francesco; Neri, Patrizia; Orlandi, Antonio; Gagliardi, Maria Scola; Velardi, Angela; Piccinetti, Corrado; Zaccanti, Francesco	Unite research with what citizens do for fun: recreational monitoring" of marine biodiversity	2010
Stafford, Richard; Hart, Adam G.; Collins, Laura; Kirkhope, Claire L.; Williams, Rachel L.; Rees, Samuel G.; Lloyd, Jane R.; Goodenough, Anne E.	Eu-Social Science: The Role of Internet Social Networks in the Collection of Bee Biodiversity Data	2010
Suzuki, Takao; Sasaki, Miki	Civil procedure for researching benthic invertebrate animals inhabiting tidal flats in eastern Japan	2010
Arvanitidis, Christos; Faulwetter, Sarah; Chatzigeorgiou, Georgios; Penev, Lyubomir; Banki, Olaf; Dailianis, Thanos; Pafilis, Evangelos; Kouratoras, Michail; Chatzinikolaou, Eva; Fanini, Lucia; Vasileiadou, Aikaterini; Pavloudi, Christina; Vavilis, Panagiotis; Koulouri, Panayota; Dounas, Costas sara bois	Engaging the broader community in biodiversity research: the concept of the COMBER pilot project for divers in ViBRANT	2011
Bramanti, Lorenzo; Vielmini, Ilaria; Rossi, Sergio; Stolfa, Stefano; Santangelo, Giovanni	Involvement of recreational scuba divers in emblematic species monitoring: The case of Mediterranean red coral (Corallium rubrum)	2011
Davies, L.; Bell, J. N. B.; Bone, J.; Head, M.; Hill, L.; Howard, C.; Hobbs, S. J.; Jones, D. T.; Power, S. A.; Rose, N.; Ryder, C.; Seed, L.; Stevens, G.; Toumi, R.; Voulvoulis, N.; White, P. C. L.	Open Air Laboratories (OPAL): A community-driven research programme	2011
De Angelo, Carlos; Paviolo, Agustin; Rode, Daniela; Cullen, Laury, Jr.; Sana, Denis; Abreu, Kaue Cachuba; da Silva, Marina Xavier; Bertrand, Anne-Sophie; Haag, Taiana; Lima, Fernando; Rinaldi, Alcides Ricieri; Fernandez, Sixto; Ramirez, Fredy; Velazquez, Myriam; Corio, Cristian; Hasson, Esteban; Di Bitetti, Mario S.	Participatory networks for large-scale monitoring of large carnivores: pumas and jaguars of the Upper Parana Atlantic Forest	2011
Kremen, C.; Ullmann, K. S.; Thorp, R. W.	Evaluating the Quality of Citizen-Scientist Data on Pollinator Communities	2011
Belt, Jami J.; Krausman, Paul R.	Evaluating Population Estimates of Mountain Goats Based on Citizen Science	2012
Cox, T. E.; Philippoff, J.; Baumgartner, E.; Smith, C. M.	Expert variability provides perspective on the strengths and weaknesses of citizen-driven intertidal monitoring program	2012
Deguines, Nicolas; Julliard, Romain; de Flores, Mathieu; Fontaine, Colin	The Whereabouts of Flower Visitors: Contrasting Land-Use Preferences Revealed by a Country-Wide Survey Based on Citizen Science	2012
Gollan, John; de Bruyn, Lisa Lobry; Reid, Nick; Wilkie, Lance	Can Volunteers Collect Data that are Comparable to Professional Scientists? A Study of Variables Used in Monitoring the Outcomes of Ecosystem Rehabilitation	2012

Tonachella, Nicolo; Nastasi, Aurora; Kaufman, Gregory; Maldini, Daniela; Rankin, Robert William	Predicting trends in humpback whale (Megaptera novaeangliae) abundance using citizen science	2012
Slade, Eleanor M.; Merckx, Thomas; Riutta, Terhi; Bebber, Daniel P.; Redhead, David; Riordan, Philip; Macdonald, David W.	Life-history traits and landscape characteristics predict macro-moth responses to forest fragmentation	2013
Bodilis, P.; Louisy, P.; Draman, M.; Arceo, H. O.; Francour, P.	Can Citizen Science Survey Non-indigenous Fish Species in the Eastern Mediterranean Sea?	2014
Buesching, Christina D.; Newman, Chris; Macdonald, David W.	How dear are deer volunteers: the efficiency of monitoring deer using teams of volunteers to conduct pellet group counts	2014
Bulleri, Fabio; Benedetti-Cecchi, Lisandro	Chasing fish and catching data: recreational spearfishing videos as a tool for assessing the structure of fish assemblages on shallow rocky reefs	2014
Casanovas, Paula; Lynch, Heather J.; Fagan, William F.	Using citizen science to estimate lichen diversity	2014
Graham, Jason R.; Tan, Qin; Jones, Linda C.; Ellis, James D.	Native Buzz: Citizen scientists creating nesting habitat for solitary bees and wasps	2014
Weinstein, Anna; Trocki, Linda; Levalley, Ron; Doster, Robert H.; Distler, Trish; Krieger, Katherine	A FIRST POPULATION ASSESSMENT OF BLACK OYSTERCATCHER HAEMATOPUS BACHMANI IN CALIFORNIA	2014
Arévalo, J.Edgardo; Méndez, Yoryineth; Roberts, Mia; Alvarado, Geiner; Vargas, Sergio	Monitoring species of mammals using track collection by rangers in the Tilarán mountain range, Costa Rica	2015
Barlow, K. E.; Briggs, P. A.; Haysom, K. A.; Hutson, A. M.; Lechiara, N. L.; Racey, P. A.; Walsh, A. L.; Langton, S. D.	Citizen science reveals trends in bat populations: The National Bat Monitoring Programme in Great Britain	2015
Biggs, Jeremy; Ewald, Naomi; Valentini, Alice; Gaboriaud, Coline; Dejean, Tony; Griffiths, Richard A.; Foster, Jim; Wilkinson, John W.; Arnell, Andy; Brotherton, Peter; Williams, Penny; Dunn, Francesca	Using eDNA to develop a national citizen science-based monitoring programme for the great crested newt (Triturus cristatus)	2015
Bosch, Stefan; Lachmann, Lars	Population trends of abundant garden birds in Baden-Wurttemberg 2005-2014: Results of the first 10 years of the citizen-science project 'Hour of the Garden Birds'.	2015
Branchini, Simone; Pensa, Francesco; Neri, Patrizia; Tonucci, Bianca Maria; Mattielli, Lisa; Collavo, Anna; Sillingardi, Maria Elena; Piccinetti, Corrado; Zaccanti, Francesco; Goffredo, Stefano	Using a citizen science program to monitor coral reef biodiversity through space and time	2015
Buldrini, Fabrizio; Simoncelli, Antinisca; Accordi, Stefania; Pezzi, Giovanna; Dallai, Daniele	Ten years of citizen science data collection of wetland plants in an urban protected area	2015
Cartwright, Lyndsay A.; Cvetkovic, Maja; Graham, Spencer; Tozer, Douglas; Chow-Fraser, Patricia	URBAN: Development of a Citizen Science Biomonitoring Program Based in Hamilton, Ontario, Canada	2015
Hawthorne, T. L.; Elmore, V.; Strong, A.; Bennett-Martin, P.; Finnie, J.; Parkman, J.; Harris, T.; Singh, J.; Edwards, L.; Reed, J.	Mapping non-native invasive species and accessibility in an urban forest: A case study of participatory mapping and citizen science in Atlanta, Georgia	2015
Kirkendale, L.; Slack-Smith, S.; Fromont, J.; Teufel, D.; Short, P.; Richards, Z.; Hosie, A.; Bryce, M.; Read, S.; Gomez, O.	RESULTS OF THE FIRST INTERTIDAL CITIZEN SCIENCE PROJECT IN WA: PORT HEDLAND COMMUNITY REEF SURVEY MARCH APRIL 2014	2015
Wilson, John-James; Jisming-See, Shi-Wei; Brandon-Mong, Guo-Jie; Lim, Aik-Hean; Lim, Voon-Ching; Lee, Ping-Shin;	Citizen Science: The First Peninsular Malaysia Butterfly Count	2015

Sing, Kong-Wah		
Abe, Jonathan; Alop-Mabuti, Aleena; Burger, Peyton; Button, Jackson; Ellsberry, Madeline; Hitzeman, Jaycinth; Morgenstern, David; Nunies, Kasey; Strother, Mara; Darling-Munson, Jared; Chan, Yvonne L.; Cassady, Robert; Vasconcellos, Sarah Maile K.; Iseman, Michael D.; Chan, Edward D.; Honda, Jennifer R.	Comparing the temporal colonization and microbial diversity of showerhead biofilms in Hawai'i and Colorado	2016
Crucitti, Pierangelo; Brocchieri, Davide; Bubbico, Francesco; Tringali, Luca; Vigliotti, Francesco	The employment of Citizen Science the study of biodiversity in a case from the Campagna Romana (Latium)	2016
Dolrenry, Stephanie; Hazzah, Leela; Frank, Laurence G.	Conservation and monitoring of a persecuted African lion population by Maasai warriors	2016
Flower, Emily; Jones, Darryl; Bernede, Lilia	Can Citizen Science Assist in Determining Koala (Phascolarctos cinereus) Presence in a Declining Population?	2016
Gerovasileiou, Vasilis; Dailianis, Thanos; Panteri, Emmanouela; Michalakis, Nikitas; Gatti, Giulia; Sini, Maria; Dimitriadis, Charalampos; Issaris, Yiannis; Salomidi, Maria; Filiopoulou, Irene; Dogan, Alper; d'Avray, Laure Thierry de Ville; David, Romain; Cinar, Melih Ertan; Koutsoubas, Drosos; Feral, Jean-Pierre; Arvanitidis, Christos	CIGESMED for divers: Establishing a citizen science initiative for the mapping and monitoring of coralligenous assemblages in the Mediterranean Sea	2016
Hardwick, Bess; Kaartinen, Riikka; Koponen, Martti; Roslin, Tomas	A rapid assessment of a poorly known insect group	2016
Katani, Josiah Z.; Mustalahti, Irmeli; Mukama, Kusaga; Zahabu, Eliakimu	Participatory forest carbon assessment in south-eastern Tanzania: experiences, costs and implications for REDD plus initiatives	2016
Ladin, Zachary S.; Higgins, Conor D.; Schmit, John Paul; Sanders, Geoffrey; Johnson, Mark J.; Weed, Aaron S.; Marshall, Matthew R.; Campbell, J. Patrick; Comiskey, James A.; Shriver, W. Gregory	Using regional bird community dynamics to evaluate ecological integrity within national parks	2016
Mahard, Tyler J.; Litvaitis, John A.; Tate, Patrick; Reed, Gregory C.; Broman, Derek J. A.	An Evaluation of Hunter Surveys to Monitor Relative Abundance of Bobcats	2016
Roelfsema, Chris; Thurstan, Ruth; Beger, Maria; Dudgeon, Christine; Loder, Jennifer; Kovacs, Eva; Gallo, Michele; Flower, Jason; Cabrera, K-le Gomez; Ortiz, Juan; Lea, Alexandra; Kleine, Diana	A Citizen Science Approach: A Detailed Ecological Assessment of Subtropical Reefs at Point Lookout, Australia	2016
Campanaro, Alessandro; Hardersen, Sonke; De Zan, Lara Redolfi; Antonini, Gloria; Bardiani, Marco; Maura, Michela; Maurizi, Emanuela; Mosconi, Fabio; Zauli, Agnese; Bologna, Marco Alberto; Roversi, Pio Federico; Peverieri, Giuseppino Sabbatini; Mason, Franco	Analyses of occurrence data of protected insect species collected by citizens in Italy	2017
Davis, Adrian; Major, Richard E.; Taylor, Charlotte E.; Martin, John M.	Novel tracking and reporting methods for studying large birds in urban landscapes	2017
Ens, E. J.; Bentley-Toon, S.; Campion, F.; Campion, S.; Kelly, J.; Towler, G.	Rapid appraisal links feral buffalo with kunkod (Melaleuca spp.) decline in freshwater billabongs of tropical northern Australia	2017
Ganey, Joseph L.; Iniguez, Jose M.;	Developing a Monitoring Program for Bird	2017

Sanderlin, Jamie S.; Block, William M.	Populations in the Chiricahua Mountains, Arizona, Using Citizen Observers: Initial Stages.	
Kallimanis, A. S.; Panitsa, M.; Dimopoulos, P.	Quality of non-expert citizen science data collected for habitat type conservation status assessment in Natura 2000 protected areas	2017
Kays, Roland; Parsons, Arielle W.; Baker, Megan C.; Kalies, Elizabeth L.; Forrester, Tavis; Costello, Robert; Rota, Christopher T.; Millspaugh, Joshua J.; McShea, William J.	Does hunting or hiking affect wildlife communities in protected areas?	2017
Long, Seh-Ling; Azmi, Nazirul A.	USING PHOTOGRAPHIC IDENTIFICATION TO MONITOR SEA TURTLE POPULATIONS AT PERHENTIAN ISLANDS MARINE PARK IN MALAYSIA	2017
Matabos, Marjolaine; Hoeberechts, Maia; Doya, Carol; Aguzzi, Jacopo; Nephin, Jessica; Reimchen, Thomas E.; Leaver, Steve; Marx, Roswitha M.; Albu, Alexandra Branzan; Fier, Ryan; Fernandez-Arcaya, Ulla; Juniper, S. Kim	Expert, Crowd, Students or Algorithm: who holds the key to deep-sea imagery 'big data' processing?	2017
Mendez, Marcos; de Jaime, Chabier; Alcantara, Manuel A.	Habitat description and interannual variation in abundance and phenology of the endangered beetle Lucanus cervus L. (Coleoptera) using citizen science monitoring	2017
Newson, Stuart E.; Evans, Hazel E.; Gillings, Simon; Jarrett, David; Raynor, Robert; Wilson, Mark W.	Large-scale citizen science improves assessment of risk posed by wind farms to bats in southern Scotland	2017
Shupe, Scott M.	High resolution stream water quality assessment in the Vancouver, British Columbia region: a citizen science study	2017
Suzuki-Ohno, Yukari; Yokoyama, Jun; Nakashizuka, Tohru; Kawata, Masakado	Utilization of photographs taken by citizens for estimating bumblebee distributions	2017
Austen, Gail E.; Bindemann, Markus; Griffiths, Richard A.; Roberts, David L.	Species identification by conservation practitioners using online images: accuracy and agreement between experts	2018
Barrows, A. P. W.; Cathey, S. E.; Petersen, C. W.	Marine environment microfiber contamination: Global patterns and the diversity of microparticle origins	2018
Campbell, Heather; Engelbrecht, Ian	The Baboon Spider Atlas - using citizen science and the fear factor' to map baboon spider (Araneae: Theraphosidae) diversity and distributions in Southern Africa	2018
Eritja, Roger; Rubido-Bara, Marga; Delacour-Estrella, Sarah; Bengoa, Mikel; Ruiz-Arrondo, Ignacio	Citizen science and biodiversity: first record of Aedes (Fredwardsius) vittatus (Bigot, 1861) (Diptera, Culicidae) in Galicia, by the means of the Mosquito Alert platform	2018
Farhadinia, Mohammad S.; Moll, Remington J.; Montgomery, Robert A.; Ashrafi, Sohrab; Johnson, Paul J.; Hunter, Luke T. B.; Macdonald, David W.	Citizen science data facilitate monitoring of rare large carnivores in remote montane landscapes	2018
Jones, Fiona M.; Allen, Campbell; Arteta, Carlos; Arthur, Joan; Black, Caitlin; Emmerson, Louise M.; Freeman, Robin; Hines, Greg; Lintott, Chris J.; Machackova, Zuzana; Miller, Grant; Simpson, Rob; Southwell, Colin; Torsey, Holly R.; Zisserman, Andrew; Hart, Tom	Time-lapse imagery and volunteer classifications from the Zooniverse Penguin Watch project	2018
Killion, Alexander K.; Roloff, Gary J.; Mayhew, Sarah; Campa, Henry, III; Winterstein, Scott	Implementing and Evaluating a Citizen-Science Program to Support Wildlife Management: MI-MAST	2018
Kortmann, Mareike; Heurich, Marco; Latifi, Hooman; Roesner, Sascha; Seidl,	Forest structure following natural disturbances and early succession provides habitat for two avian	2018

Rupert; Mueller, Joeg; Thorn, Simon	flagship species, capercaillie (Tetrao urogallus) and hazel grouse (Tetrastes bonasia)	
Marizzi, Christine; Florio, Antonia; Lee, Melissa; Khalfan, Mohammed; Ghiban, Cornel; Nash, Bruce; Dorey, Jenna; McKenzie, Sean; Mazza, Christine; Cellini, Fabiana; Baria, Carlo; Bepat, Ron; Cosentino, Lena; Dvorak, Alexander; Gacevic, Amina; Guzman-Moumtzis, Cristina; Heller, Francesca; Holt, Nicholas Alexander; Horenstein, Jeffrey; Joralemon, Vincent; Kaur, Manveer; Kaur, Tanveer; Khan, Armani; Kuppan, Jessica; Laverty, Scott; Lock, Camila; Pena, Marianne; Petrychyn, Ilona; Puthenkalam, Indu; Ram, Daval; Ramos, Arlene; Scoca, Noelle; Sin, Rachel; Gonzalez, Izabel; Thakur, Akansha; Usmanov, Husan; Han, Karen; Wu, Andy; Zhu, Tiger; Micklos, David Andrew	DNA barcoding Brooklyn (New York): A first assessment of biodiversity in Marine Park by citizen scientists	2018
Martay, B.; Pearce-Higgins, J. W.	Using data from schools to model variation in soil invertebrates across the UK: The importance of weather, climate, season and habitat	2018
Martay, Blaise; Pearce-Higgins, James W.; Harris, Sarah J.; Gillings, Simon	Monitoring landscape-scale environmental changes with citizen scientists: Twenty years of land use change in Great Britain	2018
Micaroni, Valerio; Strano, Francesca; Di Franco, Davide; Langeneck, Joachim; Gravili, Cinzia; Bertolino, Marco; Costa, Gabriele; Rindi, Fabio; Froglia, Carlo; Crocetta, Fabio; Giangrande, Adriana; Nicoletti, Luisa; Medagli, Pietro; Zuccarello, Vincenzo; Arzeni, Stefano; Bo, Marzia; Betti, Federico; Mastrototaro, Francesco; Lattanzi, Loretta; Piraino, Stefano; Boero, Ferdinando	Project Biodiversity MARE Tricase: biodiversity research, monitoring and promotion at MARE Outpost (Apulia, Italy)	2018
Rykken, Jessica J.; Farrell, Brian D.	Exploring the Microwilderness of Boston Harbor Islands National Recreation Area: Terrestrial Invertebrate All Taxa Biodiversity Inventory	2018
Singh, Priyanka; Saran, Sameer; Kumar, Dheeraj; Padalia, Hitendra; Srivastava, Ashutosh; Kumar, A. Senthil	Species Mapping Using Citizen Science Approach Through IBIN Portal: Use Case in Foothills of Himalaya	2018
Thornhill, Ian; Chautard, Alice; Loiselle, Steven	Monitoring Biological and Chemical Trends in Temperate Still Waters Using Citizen Science	2018
Baker, D. J.; Clarke, R. H.; McGeoch, M. A.	The power to detect regional declines in common bird populations using continental monitoring data	2019
Begona Garcia, Maria; Luis Silva, Jose; Tejero, Pablo; Pardo, Iker; Gomez, Daniel	Tracking the long-term dynamics of plant diversity in Northeast Spain with a network of volunteers and rangers	2019
Beirne, Christopher; Meier, Amelia C.; Mbele, Alex Ebang; Menie, Guillaume Menie; Froese, Graden; Okouyi, Joseph; Poulsen, John R.	Participatory monitoring reveals village-centered gradients of mammalian defaunation in central Africa	2019
Corazza, Carla; Baraldi, Nicola; Aldrovandi, Stefano; Mazzotti, Stefano	Biodiversity for everyone: the citizen science projects of the Museum of Natural History of Ferrara between research and collections.	2019
de Juana, Fernando; Monasterio, Yeray; Escobes, Ruth; Luis Albala, Jose; Belamendia, Gorka; de Olano, Ibon; Sebastian, Jose; Webster, Brian	The Macroheterocera (Lepidoptera) of the Salburua wetlands (Vitoria-Gasteiz, Araba/Alava, Spain): a citizen science project	2019

Drummond, Faline M.; Armstrong, Doug P.	Use of distance sampling to measure long-term changes in bird densities in a fenced wildlife sanctuary	2019
Franca, Juliana Silva; Solar, Ricardo; Hughes, Robert M.; Callisto, Marcos	Student monitoring of the ecological quality of neotropical urban streams	2019
Giovos, Ioannis; Kleitou, Periklis; Poursanidis, Dimitris; Batjakas, Ioannis; Bernardi, Giacomo; Crocetta, Fabio; Doumpas, Nikolaos; Kalogirou, Stefanos; Kampouris, Thodoros E.; Keramidas, Ioannis; Langeneck, Joachim; Maximiadi, Mary; Mitsou, Eleni; Stoilas, Vasileios-Orestis; Tiralongo, Francesco; Romanidis-Kyriakidis, Georgios; Xentidis, Nicholas-Jason; Zenetos, Argyro; Katsanevakis, Stelios	Citizen-science for monitoring marine invasions and stimulating public engagement: a case project from the eastern Mediterranean	2019
Giovos, Ioannis; Stoilas, Vasilis-Orestis; Al-Mabruk, Sara A. A.; Doumpas, Nikolaos; Marakis, Philippos; Maximiadi, Mary; Moutopoulos, Dimitrios; Kleitou, Periklis; Keramidas, Ioannis; Tiralongo, Francesco; de Maddalena, Alessandro	Integrating local ecological knowledge, citizen science and long-term historical data for endangered species conservation: Additional records of angel sharks (Chondrichthyes: Squatinidae) in the Mediterranean Sea	2019
He, Yurong; Parrish, Julia K.; Rowe, Shawn; Jones, Timothy	Evolving interest and sense of self in an environmental citizen science program	2019
Hisasue, Y.; Hisamatsu, S.; Murakami, H.	Exotic ant species found by monitoring surveys for suspected red imported fire ants, including submissions by citizens to the Biodiversity Center, Ehime Prefectural Institute of Public Health and Environmental Science during 2017.	2019
Mason, Lisa; Arathi, H. S.	Assessing the efficacy of citizen scientists monitoring native bees in urban areas	2019
Matear, Liam; Robbins, James R.; Hale, Michelle; Potts, Jonathan	Cetacean biodiversity in the Bay of Biscay: Suggestions for environmental protection derived from citizen science data	2019
Perez-Belmont, Patricia; Alvarado, Jannice; Vazquez-Salvador, Nallely; Rodriguez, Erika; Valiente, Elsa; Diaz, Julio	Water quality monitoring in the Xochimilco peri-urban wetland: experiences engaging in citizen science	2019
Pescott, Oliver L.; Walker, Kevin J.; Harris, Felicity; New, Hayley; Cheffings, Christine M.; Newton, Niki; Jitlal, Mark; Redhead, John; Smart, Simon M.; Roy, David B.	The design, launch and assessment of a new volunteer-based plant monitoring scheme for the United Kingdom	2019
Puan, Chong Leong; Yeong, Kok Loong; Ong, Kang Woei; Fauzi, Muhd Izzat Ahmad; Yahya, Muhammad Syafiq; Khoo, Swee Seng	Influence of landscape matrix on urban bird abundance: evidence from Malaysian citizen science data	2019
Qian, Haiyuan; Yu, Jianping; Shen, Xiaoli; Ding, Ping; Li, Sheng	Diversity and composition of birds in the Qianjiangyuan National Park pilot	2019
Rafiq, Kasim; Bryce, Caleb M.; Rich, Lindsey N.; Coco, Carli; Miller, David A. W.; Meloro, Carlo; Wich, Serge A.; McNutt, John W.; Hayward, Matthew W.	Tourist photographs as a scalable framework for wildlife monitoring in protected areas	2019
Robinne, Francois-Nicolas; Gallagher, Louise; Brethaut, Christian; Schlaepfer, Martin A.	A novel tool for measuring the penetration of the ecosystem service concept into public policy	2019

Schade, Sven; Kotsev, Alexander; Cardoso, Ana Cristina; Tsiamis, Konstantinos; Gervasini, Eugenio; Spinelli, Fabiano; Mitton, Irena; Sgnaolin, Roberto	Aliens in Europe. An open approach to involve more people in invasive species detection	2019
Schuttler, Stephanie G.; Sears, Rebecca S.; Orendain, Isabel; Khot, Rahul; Rubenstein, Daniel; Rubenstein, Nancy; Dunn, Robert R.; Baird, Elizabeth; Kandros, Kimberly; O'Brien, Timothy; Kays, Roland	Citizen Science in Schools: Students Collect Valuable Mammal Data for Science, Conservation, and Community Engagement	2019
Smale, Dan A.; Epstein, Graham; Parry, Mark; Attrill, Martin J.	Spatiotemporal variability in the structure of seagrass meadows and associated macrofaunal assemblages in southwest England (UK): Using citizen science to benchmark ecological pattern	2019
Sumner, Seirian; Bevan, Peggy; Hart, Adam G.; Isaac, Nicholas J. B.	Mapping species distributions in 2 weeks using citizen science	2019
Yardi, Kranti D.; Bharucha, Erach; Girade, Swapnil	Post-restoration monitoring of water quality and avifaunal diversity of Pashan Lake, Pune, India using a citizen science approach	2019
Appenfeller, Logan R.; Lloyd, Sarah; Szendrei, Zsofia	Citizen science improves our understanding of the impact of soil management on wild pollinator abundance in agroecosystems	2020
Blake, Charlie; Rhanor, Allison K.	The impact of channelization on macroinvertebrate bioindicators in small order Illinois streams: insights from long-term citizen science research	2020
Bonnet-Lebrun, A-S; Karamanlidis, A. A.; de Gabriel Hernando, M.; Renner, I; Gimenez, O.	Identifying priority conservation areas for a recovering brown bear population in Greece using citizen science data	2020
Castracani, Cristina; Spotti, Fiorenza Augusta; Schifani, Enrico; Giannetti, Daniele; Ghizzoni, Martina; Grasso, Donato Antonio; Mori, Alessandra	Public Engagement Provides First Insights on Po Plain Ant Communities and Reveals the Ubiquity of the Cryptic Species Tetramorium immigrans (Hymenoptera, Formicidae)	2020
Deguines, Nicolas; Prince, Karine; Prevot, Anne-Caroline; Fontaine, Benoit	Assessing the emergence of pro-biodiversity practices in citizen scientists of a backyard butterfly survey	2020
Ebihaha, Kengo; Yasukawa, Masaki; Nagai, Mihoko; Kitsuregawa, Masaru; Washitani, Izumi	Feasibility of citizen science monitoring of mutualistic networks between butterflies and plants in Tokyo, Japan.	2020
Edgar, Graham J.; Cooper, Antonia; Baker, Susan C.; Barker, William; Barrett, Neville S.; Becerro, Mikel A.; Bates, Amanda E.; Brock, Danny; Ceccarelli, Daniela M.; Clausius, Ella; Davey, Marlene; Davis, Tom R.; Day, Paul B.; Green, Andrew; Griffiths, Samuel R.; Hicks, Jamie; Hinojosa, Ivan A.; Jones, Ben K.; Kininmonth, Stuart; Larkin, Meryl F.; Lazzari, Natali; Lefcheck, Jonathan S.; Ling, Scott D.; Mooney, Peter; Oh, Elizabeth; Perez-Matus, Alejandro; Pocklington, Jacqueline B.; Riera, Rodrigo; Sanabria-Fernandez, Jose A.; Seroussi, Yanir; Shaw, Ian; Shields, Derek; Shields, Joe; Smith, Margo; Soler, German A.; Stuart-Smith, Jemina; Turnbull, John; Stuart-Smith, Rick D.	Reef Life Survey: Establishing the ecological basis for conservation of shallow marine life	2020
Gardiner, Tim; Didham, Raphael K.	Glowing, glowing, gone? Monitoring long-term trends in glow-worm numbers in south-east England	2020
Gili, Fabrizio; Newson, Stuart E.; Gillings, Simon; Chamberlain, Dan E.; Border, Jennifier A.	Bats in urbanising landscapes: habitat selection and recommendations for a sustainable future	2020

Gizzi, Francesca; Jimenez, Jesus; Schaefer, Susanne; Castro, Nuno; Costa, Sonia; Lourenco, Silvia; Jose, Ricardo; Canning-Clode, Joao; Monteiro, Joao	Before and after a disease outbreak: Tracking a keystone species recovery from a mass mortality event	2020
Lanner, Julia; Huchler, Katharina; Pachinger, Baerbel; Sedivy, Claudio; Meimberg, Harald	Dispersal patterns of an introduced wild bee, Megachile sculpturalis Smith, 1853 (Hymenoptera: Megachilidae) in European alpine countries	2020
Monterastelli, Elisa; Poloni, Riccardo	The Insetti.A.MO project: an insect population census in Modena city (Italy)	2020
Nunes, Miguel Simoes; Falconer, Kristie; Jelic, Dusan; Martin, Thomas Edward; Kucinic, Mladen; Jocque, Merlijn	The value of eco-volunteer projects for biodiversity conservation: butterfly monitoring in Krka National Park (Croatia) with an updated checklist	2020
Platenberg, Renata J.; Raymore, Martha; Primack, Avram; Troutman, Kelcie	Monitoring Vocalizing Species by Engaging Community Volunteers Using Cell Phones	2020
Rameli, Nurul I. A. Mohd; Lappan, Susan; Bartlett, Thad Q.; Ahmad, Siti K.; Ruppert, Nadine	Are social media reports useful for assessing small ape occurrence? A pilot study from Peninsular Malaysia	2020
Schneiderhan-Opel, Jennifer; Bogner, Franz X.	How fascination for biology is associated with students' learning in a biodiversity citizen science project	2020
Shah, Md Nur Ahad; Khan, Md Kawsar	OdoBD: An online database for the dragonflies and damselflies of Bangladesh	2020
Shang Xiaotong; Luo Chunping; Li Bin; Zheng Yong; Zhou Zhiqiang; Zhang Li; Li Sheng	Diversity and Fauna Composition of Birds in the Wanglang National Nature Reserve, Sichuan.	2020
Sheard, Julie K.; Sanders, Nathan J.; Gundlach, Carsten; Schar, Sami; Larsen, Rasmus Stenbak	Monitoring the influx of new species through citizen science: the first introduced ant in Denmark	2020
Stenhouse, Alan; Roetman, Philip; Lewis, Megan; Koh, Lian Pin	Koala Counter: Recording Citizen Scientists' search paths to Improve Data Quality	2020
Tiralongo, Francesco; Crocetta, Fabio; Riginella, Emilio; Lillo, Antonio Oscar; Tondo, Elena; Macali, Armando; Mancini, Emanuele; Russo, Fabio; Coco, Salvatore; Paolillo, Giuseppe; Azzurro, Ernesto	Snapshot of rare, exotic and overlooked fish species in the Italian seas: A citizen science survey	2020
Uhrin, Amy V.; Lippiatt, Sherry; Herring, Carlie E.; Dettloff, Kyle; Bimrose, Kate; Butler-Minor, Chris	Temporal Trends and Potential Drivers of Stranded Marine Debris on Beaches Within Two US National Marine Sanctuaries Using Citizen Science Data	2020
Werenkraut, Victoria; Baudino, Florencia; Roy, Helen E.	Citizen science reveals the distribution of the invasive harlequin ladybird (Harmonia axyridisPallas) in Argentina	2020
Wotton, S. R.; Eaton, M. A.; Sheehan, D.; Munyekenye, F. Barasa; Burfield, I. J.; Butchart, S. H. M.; Moleofi, K.; Nalwanga-Wabwire, D.; Ndang'ang'a, P. K.; Pomeroy, D.; Senyatso, K. J.; Gregory, R. D.	Developing biodiversity indicators for African birds	2020
Ahmad, Abrar; Gary, Demi; Rodiansyah; Sinta; Srifitria; Putra, Wahyu; Sagita, Novia; Adirahmanta, Sadtata Noor; Miller, Adam E.	Leveraging local knowledge to estimate wildlife densities in bornean tropical rainforests	2021
Alther, Roman; Bongni, Nicole; Borko, Spela; Fiser, Cene; Altermatt, Florian	Citizen science approach reveals groundwater fauna in Switzerland and a new species of Niphargus (Amphipoda, Niphargidae)	2021
Anton, Victor; Germishuys, Jannes; Bergstrom, Per; Lindegarth, Mats; Obst, Matthias	An open-source, citizen science and machine learning approach to analyse subsea movies	2021
Arbelaez-Cortes, Enrique; Sanchez-	EXPERIENCES OF SURVEYING URBAN BIRDS DURING	2021

Sarria, Camilo E.; Ocampo, David; Estela, Felipe A.; Garcia-Arroyo, Michelle; MacGregor-Fors, Ian	THE ANTHROPAUSE IN COLOMBIA	
Aura, Christopher Mulanda; Nyamweya, Chrisphine S.; Owiti, Horace; Odoli, Cyprian; Musa, Safina; Njiru, James M.; Nyakeya, Kobingi; Masese, Frank O.	Citizen Science for Bio-indication: Development of a Community-Based Index of Ecosystem Integrity for Assessing the Status of Afrotropical Riverine Ecosystems	2021
Balciauskas, Linas; Balciauskiene, Laima; Litvaitis, John A.; Tijusas, Eugenijus	Adaptive monitoring: using citizen scientists to track wolf populations when winter-track counts become unreliable	2021
Biddle, Rebecca; Solis-Ponce, Ivette; Jones, Martin; Marsden, Stuart; Pilgrim, Mark; Devenish, Christian	The value of local community knowledge in species distribution modelling for a threatened Neotropical parrot	2021
Encarnacao, Joao; Baptista, Vania; Teodosio, Maria Alexandra; Morais, Pedro	Low-Cost Citizen Science Effectively Monitors the Rapid Expansion of a Marine Invasive Species	2021
Flaminio, Simone; Ranalli, Rosa; Zavatta, Laura; Galloni, Marta; Bortolotti, Laura	Beewatching: A Project for Monitoring Bees through Photos	2021
Gadsden, Gabriel, I; Malhotra, Rumaan; Schell, Justin; Carey, Tiffany; Harris, Nyeema C.	Michigan ZoomIN: Validating Crowd-Sourcing to Identify Mammals from Camera Surveys	2021
Garcia, Maria B.; Silva, Jose L.; Tejero, Pablo; Pardo, Iker	Detecting early-warning signals of concern in plant populations with a Citizen Science network. Are threatened and other priority species for conservation performing worse?	2021
Gutierrez-Munoz, Paula; Walters, Alice E. M.; Dolman, Sarah J.; Pierce, Graham J.	Patterns and Trends in Cetacean Occurrence Revealed by Shorewatch, a Land-Based Citizen Science Program in Scotland (United Kingdom)	2021
Kalaentzis, Konstantinos; Kazilas, Christos; Demetriou, Jakovos; Koutsoukos, Evangelos; Avtzis, Dimitrios N.; Georgiadis, Christos	Alientoma, a Dynamic Database for Alien Insects in Greece and Its Use by Citizen Scientists in Mapping Alien Species	2021
Kasten, Paula; Jenkins, Stuart R.; Christofoletti, Ronaldo A.	Participatory Monitoring-A Citizen Science Approach for Coastal Environments	2021
Kirchhoff, Casey; Callaghan, Corey T.; Keith, David A.; Indiarto, Dony; Taseski, Guy; Ooi, Mark Kj; Le Breton, Tom D.; Mesaglio, Thomas; Kingsford, Richard T.; Cornwell, William K.	Rapidly mapping fire effects on biodiversity at a large-scale using citizen science	2021
Lee, Tracy S.; Kahal, Nicole L.; Kinas, Holly L.; Randall, Lea A.; Baker, Tyne M.; Carney, Vanessa A.; Kendell, Kris; Sanderson, Ken; Duke, Danah	Advancing Amphibian Conservation through Citizen Science in Urban Municipalities	2021
Lin, Meixi; Simons, Ariel Levi; Harrigan, Ryan J.; Curd, Emily E.; Schneider, Fabian D.; Ruiz-Ramos, Dannise V.; Gold, Zack; Osborne, Melisa G.; Shirazi, Sabrina; Schweizer, Teia M.; Moore, Tiara N.; Fox, Emma A.; Turba, Rachel; Garcia-Vedrenne, Ana E.; Helman, Sarah K.; Rutledge, Kelsi; Mejia, Maura Palacios; Marwayana, Onny; Munguia Ramos, Miroslava N.; Wetzer, Regina; Pentcheff, N. Dean; McTavish, Emily Jane; Dawson, Michael N.; Shapiro, Beth; Wayne, Robert K.; Meyer, Rachel S.	Landscape analyses using eDNA metabarcoding and Earth observation predict community biodiversity in California	2021
Machado, Augusto A.; Bertoncini, Athila A.; Santos, Luciano N.; Creed, Joel C.; Masi, Bruno P.	Participatory monitoring of marine biological invaders: a novel program to include citizen scientists	2021
Mangelli, Tarcio S.; Zapelini, Cleverson;	Voluntary scuba diving as a method for monitoring	2021

da Rocha, Wesley Duarte; Schiavetti, Alexandre	invasive exotic marine species	
Mesaglio, Thomas; Soh, Aaron; Kurniawidjaja, Steven; Sexton, Chuck	'First Known Photographs of Living Specimens': the power of iNaturalist for recording rare tropical butterflies	2021
Meschini, Marta; Machado Toffolo, Mariana; Marchini, Chiara; Caroselli, Erik; Prada, Fiorella; Mancuso, Arianna; Franzellitti, Silvia; Locci, Laura; Davoli, Marco; Trittoni, Michele; Nanetti, Enrico; Tittarelli, Mara; Bentivogli, Riccardo; Branchini, Simone; Neri, Patrizia; Goffredo, Stefano	Reliability of Data Collected by Volunteers: A Nine-Year Citizen Science Study in the Red Sea	2021
Meyer, Rachel S.; Ramos, Miroslava Munguia; Lin, Meixi; Schweizer, Teia M.; Gold, Zachary; Ramos, Dannise Ruiz; Shirazi, Sabrina; Kandlikar, Gaurav; Kwan, Wai-Yin; Curd, Emily E.; Freise, Amanda; Parker, Jordan Moberg; Sexton, Jason P.; Wetzer, Regina; Pentcheff, N. Dean; Wall, Adam R.; Pipes, Lenore; Garcia-Vedrenne, Ana; Mejia, Maura Palacios; Moore, Tiara; Orland, Chloe; Ballare, Kimberly M.; Worth, Anna; Beraut, Eric; Aronson, Emma L.; Nielsen, Rasmus; Lewin, Harris A.; Barber, Paul H.; Wall, Jeff; Kraft, Nathan; Shapiro, Beth; Wayne, Robert K.	The CALeDNA program: Citizen scientists and researchers inventory California's biodiversity	2021
Moller, Anders Pape; Czeszczewik, Dorota; Erritzoe, Johannes; Flensted-Jensen, Einar; Laursen, Karsten; Liang, Wei; Walankiewicz, Wieslaw	Citizen Science for Quantification of Insect Abundance on Windshields of Cars Across Two Continents	2021
Moro, Arrigo; Beaurepaire, Alexis; Dall'Olio, Raffaele; Rogenstein, Steve; Blacquiere, Tjeerd; Dahle, Bjorn; de Miranda, Joachim R.; Dietemann, Vincent; Locke, Barbara; Licon Luna, Rosa Maria; Le Conte, Yves; Neumann, Peter	Using Citizen Science to Scout Honey Bee Colonies That Naturally Survive Varroa destructor Infestations	2021
Pellicioli, Luca; Cimberio, Patrizia	Citizen science project on Alpine ibex, Capra ibex, in the Orobie Alps.	2021
Rodhouse, Thomas J.; Rose, Sara; Hawkins, Trent; Rodriguez, Rogelio M.	Audible bats provide opportunities for citizen scientists	2021
Rowe, Helen, I; Gruber, Daniel; Fastiggi, Mary	Where to start? A new citizen science, remote sensing approach to map recreational disturbance and other degraded areas for restoration planning	2021
Sanden, Taru; Wawra, Anna; Berthold, Helene; Miloczki, Julia; Schweinzer, Agnes; Gschmeidler, Brigitte; Spiegel, Heide; Debeljak, Marko; Trajanov, Aneta	TeaTime4Schools: Using Data Mining Techniques to Model Litter Decomposition in Austrian Urban School Soils	2021
Squires, Thomas M.; Yuda, Pramana; Akbar, Panji Gusti; Collar, Nigel J.; Devenish, Christian; Taufiqurrahman, Imam; Wibowo, Waskito Kukuh; Winarni, Nurul L.; Yanuar, Ahmad; Marsden, Stuart J.	Citizen science rapidly delivers extensive distribution data for birds in a key tropical biodiversity area	2021
Stenhouse, Alan; Perry, Tahlia; Grutzner, Frank; Lewis, Megan; Koh, Lian Pin	EchidnaCSI - Improving monitoring of a cryptic species at continental scale using Citizen Science	2021
Sun, Catherine C.; Hurst, Jeremy E.; Fuller, Angela K.	Citizen Science Data Collection for Integrated Wildlife Population Analyses	2021

Thomaes, Arno; Barbalat, Sylvie; Bardiani, Marco; Bower, Laura; Campanaro, Alessandro; Fanega Sleziak, Natalia; Goncalo Soutinho, Joao; Govaert, Sanne; Harvey, Deborah; Hawes, Colin; Kadej, Marcin; Mendez, Marcos; Meriguet, Bruno; Rink, Markus; Rossi De Gasperis, Sarah; Ruyts, Sanne; Jelaska, Lucija Seric; Smit, John; Smolis, Adrian; Snegin, Eduard; Tagliani, Arianna; Vrezec, Al	The European Stag Beetle (Lucanus cervus) Monitoring Network: International Citizen Science Cooperation Reveals Regional Differences in Phenology and Temperature Response	2021
Townsend, Philip A.; Clare, John D. J.; Liu, Nanfeng; Stenglein, Jennifer L.; Anhalt-Depies, Christine; Van Deelen, Timothy R.; Gilbert, Neil A.; Singh, Aditya; Martin, Karl J.; Zuckerberg, Benjamin webster	Snapshot Wisconsin: networking community scientists and remote sensing to improve ecological monitoring and management	2021
Yang, Jun; Xing, Danqi; Luo, Xiangyu	Assessing the performance of a citizen science project for monitoring urban woody plant species diversity in China	2021

Appendix B-3. RAW data for openness scores per principle: -1 (closed), 0.5 (partially open), 1 (open). Blanks are found where the principle was not applicable (NA).

Openness Scores Per Principle					
DMP	*Preregistration*	*Data*	*Code*	*Software*	*Access*
0.5	1	1		1	-1
-1	-1	1	1		1
-1	-1	-1	-1	1	-1
-1	0.5	1	0.5	-1	-1
-1	-1	1	-1	-1	-1
-1	-1	-1	-1	1	-1
1	-1	1			1
-1	0.5	0.5		0.5	1
1	-1	1			1
0.5	1	0.5			1
1		1	0.5	0.5	1
-1	-1	0.5		-1	1
0.5	-1	1	-1	1	1
0.5	0.5	0.5			1
0.5	-1	1	1	1	1
-1	-1	0.5			-1
-1	0.5	1		0.5	1
0.5	-1	1	1	1	1
0.5	-1	0.5	0.5	1	1
	0.5	1	1	1	1
-1	1	1	1	0.5	-1
0.5	-1	1	1	1	1
-1	-1	1	0.5	1	-1
0.5	0.5	1		-1	1
-1	-1	1	1	1	-1
-1	-1	0.5	0.5	1	-1
-1	0.5	1			-1
-1	-1	1		1	1
1	0.5	1			1
-1	-1	1	1	0.5	1
0.5	-1	1	-1	1	1
-1	0.5	-1	-1	0.5	1
1	-1	1	1	1	1
-1	-1	-1	-1	-1	-1
-1	-1	0.5	0.5	0.5	1
-1	-1	1	-1	1	1
-1	-1	1			-1
-1	0.5	1		0.5	1
-1	-1	-1		0.5	-1
-1	0.5	1	0.5	1	1
-1	-1	1	0.5	1	1
-1	-1	1	-1	1	-1

Appendix B-4. Data for the number of citizen science projects per year and associated average annual openness score.

Year	Number of Projects	Average Annual Open Score
2005	1	0.5
2010	1	0.2
2011	1	-0.67
2012	1	-0.17
2015	3	-0.39
2016	4	0.55
2017	6	0.18
2018	2	0.5
2019	4	0.35
2020	8	0.22
2021	11	-0.05

Appendix C

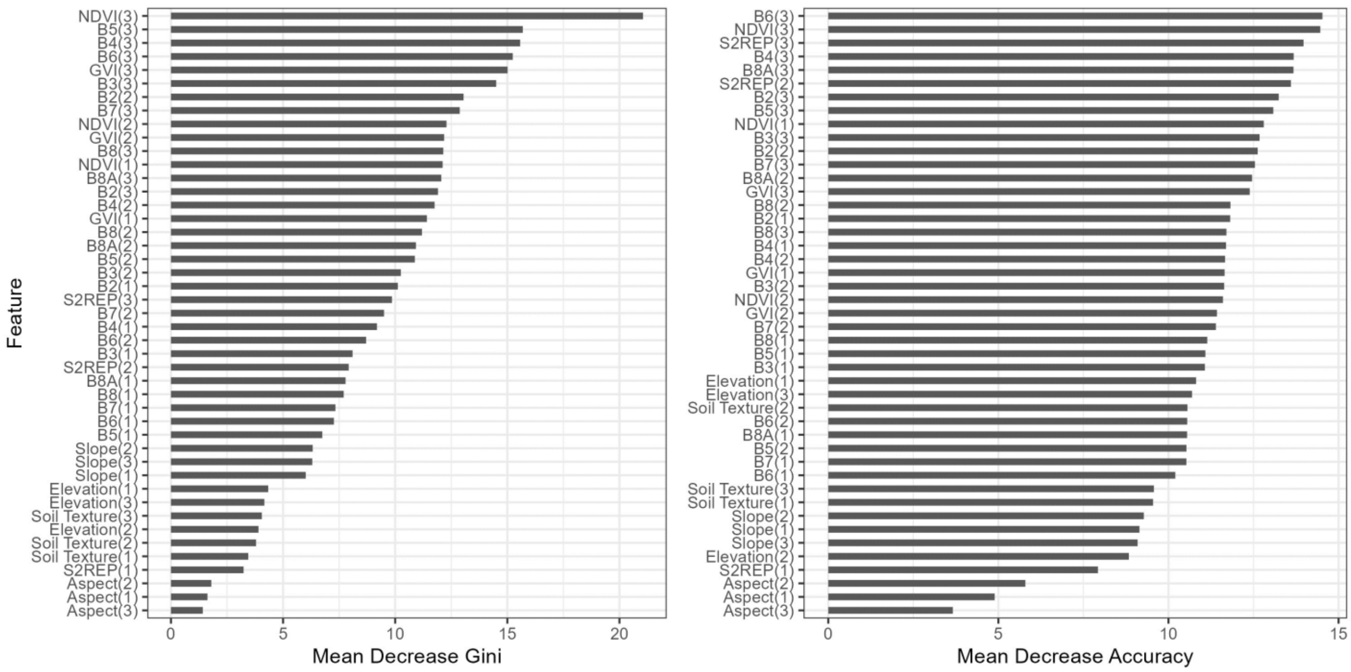
Appendix C-1. Data sources for spatial analysis of grassland distribution

Data	Source	Citation
Light Gray Canvas Base Map	ArcGIS Pro	Esri, DeLorme, HERE, MapmyIndia
UK CEH Land Cover Map	EDINA Digimap	Land Cover map of Great Britain., 2019. [TIFF geospatial data], Scale 1:250000, Tiles: GB, Updated: 30 June 2020, CEH, Using: EDINA Environment Digimap Service, <https://digimap.edina.ac.uk>, Downloaded: 2020-12-07 15:18:08.866)
HabMoS Map	NatureScot/ Scottish Natural Heritage	Contains SNH information licensed under the Open Government Licence v3.0.
HadUK-Grid Gridded Climate Observations on a 1km grid over the UK, v1.0.2.1 (1862-2019).	Met Office via CEDA	Met Office; Hollis, D.; McCarthy, M.; Kendon, M.; Legg, T.; Simpson, I. (2020): HadUK-Grid Gridded Climate Observations on a 1km grid over the UK, v1.0.2.1 (1862-2019). Centre for Environmental Data Analysis, *21 October 2020*. doi:10.5285/89908dfcb97b4a28976df806b4818639. http://dx.doi.org/10.5285/89908dfcb97b4a28976df806b4818639
GIS_SNH_OWNER.LANDSCAPE_MAP	NatureScot/ Scottish Natural Heritage	Available under the Open Government Licence (http://www.nationalarchives.gov.uk/doc/open-government-licence)
Lowland Grasslands Database	Scottish Natural Heritage	
UK BMS 2019 Site Locations	UK Butterfly Monitoring Scheme	Botham, M.; Brereton, T.; Harris, S.; Harrower, C.; Middlebrook, I.; Randle, Z.; Roy, D.B. (2020). United Kingdom Butterfly Monitoring Scheme: site location data 2019. NERC Environmental Information Data Centre. (Dataset). https://doi.org/10.5285/8a41e1c8-3018-44f1-8d0a-c1b1ad957fc9
UK BMS 2019 Site Locations with Habitat Type	UK Butterfly Monitoring Scheme/ Butterfly Conservation	Botham, M.; Brereton, T.; Harris, S.; Harrower, C.; Middlebrook, I.; Randle, Z.; Roy, D.B. (2020). United Kingdom Butterfly Monitoring Scheme: site location data 2019. NERC Environmental Information Data Centre. (Dataset). https://doi.org/10.5285/8a41e1c8-3018-44f1-8d0a-c1b1ad957fc9
Northern Brown Argus Site Locations	Butterfly Conservation	Butterfly Conservation Scottish Borders NBA survey dataset, 2021.

Appendix C-2. Sentinel-2 acquisition dates compared to in situ data survey dates across the three 2021 field campaigns.

Site	**C1**		**C2**		**C3**	
	Survey	*Sentinel-2*	*Survey*	*Sentinel-2*	*Survey*	*Sentinel-2*
AC	25/05/2021	29/05/2021	23/07/2021	01/07/2021	20/08/2021	25/08/2021
BL	28/05/2021	30/05/2021	20/07/2021	21/07/2021	17/08/2021	25/08/2021
CL	13/05/2021	07/05/2021	09/07/2021	14/07/2021	01/09/2021	28/08/2021
EH	12/05/2021	19/05/2021	05/07/2021	01/07/2021	24/08/2021	25/08/2021
GD	24/05/2021	02/05/2021	07/07/2021	01/07/2021	24/08/2021	10/08/2021
GF	27/05/2021	05/05/2021	21/07/2021	26/07/2021	18/08/2021	18/08/2021
GH	19/05/2021	27/05/2021	08/07/2021	06/07/2021	31/08/2021	28/08/2021
GL	26/05/2021	29/05/2021	22/07/2021	01/07/2021	19/08/2021	27/08/2021
HM	29/05/2021	27/05/2021	20/07/2021	21/07/2021	17/08/2021	25/08/2021
HR	25/05/2021	02/05/2021	23/07/2021	01/07/2021	20/08/2021	27/08/2021
LM	07/05/2021 +18/05/2021	24/05/2021	02/07/2021	01/07/2021	23/08/2021	25/08/2021
MM	12/05/2021 +18/05/2021	19/05/2021	02/07/2021	01/07/2021	23/08/2021	25/08/2021
MP	29/05/2021	07/05/2021	19/07/2021	21/07/2021	16/08/2021	25/08/2021
RP	14/05/2021	27/05/2021	19/07/2021	19/07/2021	16/08/2021	25/08/2021
SA	24/05/2021	22/05/2021	24/07/2021	23/07/2021	21/08/2021	01/09/2021
SM	11/05/2021	02/05/2021	06/07/2021	01/07/2021	26/08/2021	25/08/2021

Appendix C-3. Feature importance of the predictor variables used in the Random Forest classification model of species-rich grasslands. Features are numbered by season: 1 = early season (May), 2 = mid-season (June/July), 3 = end of season (August).



Appendix D – Chapter 4 Supplementary Materials 1

Appendix D-1. Study site names with their corresponding codes.

Site Name	Site Code
Auchtermuchty Common	AC
Cleugh	CL
Eildon Hill	EH
Greenlaw Dean	GD
Glen Fender	GF
Havoc Meadow	HM
Lindean Moor	LM
Murder Moss	MM
Muirsheil Park	MP
St Abb’s	SA
Smardale Gill	SM

Appendix D-2. Investigation of predictive power of Sentinel-2 data across the 11 study sites, plus an additional five sites surveyed in 2021, across 2021, 2022, and both years combined.

Site	Year	*n*	S2					
			AGB (g/m^2)		Sward Height (cm)		SPAD-measured chlorophyll-proxy	
			R^2	*RMSE*	R^2	*RMSE*	R^2	*RMSE*
AC	2021	18	0.289	27.35	0.299	15.77		
AC	2022	18	0.282	38.67	0.406	12.62	0.185	10.52
AC	both	36	0.117	41.91	0.503	9.07		
BL	2021	18	0.655	18.72	0.706	8.00		
CL	2021	18	0.832	14.92	0.823	12.19		
CL	2022	18	0.929	23.37	0.924	6.65	0.027	6.62
CL	both	36	0.318	37.93	0.741	8.38		
EH	2021	18	0.490	48.00	0.586	26.67		
EH	2022	18	0.925	28.49	0.622	14.97	0.655	5.98
EH	both	36	0.168	37.94	0.051	21.83		
GF	2021	18	0.094	130.18	0.290	34.83		
GF	2022	18	0.284	34.02	0.022	27.89	0.789	31.06
GF	both	36	0.156	41.03	0.013	39.80		
GL	2021	18	0.137	49.81	0.568	13.82		
GD	2021	17	0.402	37.73	0.879	7.62		
GD	2022	18	0.0001	36.67	0.399	12.99	0.242	3.14
GD	both	35	0.347	36.32	0.390	14.30		
GH	2021	18	0.429	25.03	0.380	20.13		
HM	2021	23	0.669	45.61	0.261	18.61		
HM	2022	18	0.226	140.08	0.883	9.24	0.922	7.60
HM	both	41	0.057	77.24	0.205	26.14		
HP	2021	18	0.549	42.49	0.280	21.42		
LM	2021	18	0.938	19.38	0.357	9.35		
LM	2022	18	0.856	16.19	0.776	11.80	0.304	7.01
LM	both	36	0.118	32.92	0.878	10.11		
MP	2021	18	0.219	30.82	0.110	28.22		
MP	2022	15	0.676	32.56	0.901	19.34		
MP	both	33	0.313	34.98	0.224	23.15		
MM	2021	17	0.606	30.32	0.396	24.05		
MM	2022	18	0.745	69.50	0.863	12.45	0.699	7.06
MM	both	35	0.500	31.77	0.465	11.31		
RP	2021	18	0.682	40.45	0.922	10.96		

SM	2021	17	0.326	19.83	0.121	21.36		
SM	2022	18	0.962	17.36	0.195	21.02	0.441	5.13
SM	both	35	0.581	16.00	0.693	8.66		
SA	2021	18	0.005	26.59	0.886	3.93		
SA	2022	18	0.448	18.41	0.848	4.82	0.710	5.03
SA	both	36	0.111	23.58	0.102	6.59		

Appendix E

Appendix E-1. Online resources found on the citizen science survey platform.

Identification Guides and Help

Here you will find resources that will help with your Identification skills. These range from online PDF guides or keys to free apps or guides to buy for in field. Everything here is optional and only to enhance your ID skills or help whilst out surveying. Below are the recommended resources with further suggestions at the end if you can't get enough!

Grasslands

- http://www.magnificentmeadows.org.uk/assets/pdfs/Meadow_ID_Leaflet.pdf
- https://www.speciesrecoverytrust.org.uk/resources
- https://www.wildlifetrusts.org/habitats/grassland

Plants

- https://www.npms.org.uk/sites/default/files/PDF/NPMS%20ID%20GUIDE_WEB_0.pdf
- https://www.discoverthewild.co.uk/resources
- https://gobotany.nativeplanttrust.org/simple/non-monocots/

Butterflies and Moths

- https://butterfly-conservation.org/butterflies/identify-a-butterfly
- https://www.wildlifetrusts.org/wildlife/identify-british-butterflies
- https://www.discoverthewild.co.uk/_files/ugd/562348_a3c49ceed9934331bf796426c143552d.pdf
- https://www.discoverthewild.co.uk/_files/ugd/562348_98be87b913894f4fada3025e1b3c9306.pdf

Apps (freely available on Android and Iphone):

- PlantNet
- iRecord Butterflies
- iNaturalist

Field Guides to Purchase

- https://www.field-studies-council.org/product-category/publications/?fwp_publication_type=fold-out-guide&fwp_natural_history_courses=botany%2Cbutterflies-and-moths%2Cgrasses-sedges-and-rushes%2Chabitat-surveying
- https://www.nhbs.com/a-field-guide-to-grasses-sedges-and-rushes-book?bkfno=228946&ca_id=1495&adlocale=uk&gclid=EAIaIQobChMIzIjtmc-T9wIVAuDtCh26RwlNEAQYAiABEgKOKfD_BwE

Further Help:

- https://www.woodlandtrust.org.uk/trees-woods-and-wildlife/habitats/grassland/
- http://www.brerc.eclipse.co.uk/files/brerc_grass_key.pdf

- https://www.google.co.uk/shopping/product/1?q=grasses+field+guide&prds=epd:11757374092611632455,eto:11757374092611632455_0,pid:11757374092611632455&sa=X&ved=0ahUKEwjXpYCnz5P3AhWREMAKHaavCREQ9pwGCAU
- https://www.amazon.co.uk/Wild-Flower-Key-Revised-identify/dp/0723251754/ref=asc_df_0723251754/?tag=googshopuk-21&linkCode=df0&hvadid=310872601819&hvpos=&hvnetw=g&hvrand=15841409056965496743&hvpone=&hvptwo=&hvqmt=&hvdev=c&hvdvcmdl=&hvlocint=&hvlocphy=1006605&hvtargid=pla-350354445358&psc=1&th=1&psc=1

Appendix E-2. Excerpt of grassland classification key indicator guide on the online platform for helping participants determine the grassland classification.

Grassland Classification Key Indicators

Introduction

Grasslands are diverse habitats making them quite difficult to differentiate. Due to their often transitional and mosaiced nature, it becomes tricky to fully separate the grassland classes.

Many plant and butterfly species can also tolerate a range of conditions meaning they may not be exclusive to one distinct grassland habitat. For example, Sweet Vernal grass is often found in all grassland classes.

However, key indicator species and their frequencies can help us determine what grassland class a habitat most likely belongs to.

Use these descriptors to help you assign grasslands classes to a site, basing the justification on the presence and frequency of the species found.

Acid Grasslands

- Acid grasslands, defined by their lower pH (<5) are widespread across the UK.
- Usually more nutrient poor than other classes of grasslands, with low to high species richness.
- Fine-leaved grasses are present, with mosses, and dwarf-shrubs.
- Acid grasslands are often in mosaic with heathlands and rush pastures.

https://www.speciesrecoverytrust.org.uk/_files/ugd/59de27_39c178f92ce84bae8dd349b45be0c581.pdf

Plant Indicators	Butterfly Indicators
Wavy-hair grass	Small heath
Sheep's Fescue grass	Meadow brown
Mat grass	Small copper
Deer Grass	Skippers
Purple moor grass	
Heath bedstraw	
Common tormentil	
Heather	
Bilberry	
Mosses more abundant	

Things to Note

- A lot of these species will be found in multiple grassland habitats, which may be confusing. Make sure you are looking at what is dominant, frequent, and in combination with each other
- It may also be difficult to differentiate species within the same family. If you are unsure, identify it to the highest level you can and use help of field guides and apps.
- Species often have multiple common names as well, don't let this confuse you, put down what you know…
- If you are unsure of anything do not pick or touch it. It may be rare, endangered…or poisonous!

Appendix E-3. Grassland classification data entry form.

Grassland Classification Form

Site Name:

OS Grid Ref/ Lat Long:

Nearest Town to Site:

Weather Conditions:

Date:

Age range (<18, 18-24, 25-34, 35-44, 45-54, 55-64, >64):

Previous Survey Experience and Species ID Knowledge (none/moderate/advanced):

Transect Survey: This stage of the survey will enable us to determine the dominant habitat based on what you can see while walking the site, as well as helping us to record the use of the habitat by key butterfly species. This can be done in 3 simple steps.

Firstly, either on arrival or before arrival if using Google Earth, decide on a path of roughly 1 km that you can walk with ease/safely. Secondly, while walking your transect path, you will use the key indicator species guide to identify the main grassland classification (Note this in Box 1 below). We will ask you to tell us if you have seen any common rock-rose on the site along your transect (Note this in box 5). Finally, as you walk, note down any butterfly species you observe whilst walking, in box 6.

1. Grassland Class Observed:

Acid ☐ Marshy ☐

Neutral ☐ Other ☐

Calcareous ☐

2. How confident are you in your answer, please circle score (1 = no confidence to 5 = very confident)?

1 2 3 4 5

3. If you selected "other" in question 1, please indicate the habitat type:

4. Any extra information (site condition/management practices/site threats to habitats):

Quadrat Surveys: Your next task is to give more detail on the plant species at your site. This will be done by focusing on 3 [0.3 m^2] quadrats (squares of ground, use the length of an A4 page) by telling us about the maximum height of the tallest plant (known as sward height, use the printed ruler on final page), the species richness (the number of different types of plants – don't worry, you don't need to know all their names!), the names of any species that you can identify, and their frequency on the following scale: D - Dominant > 75% cover; A - Abundant 51 - 75% cover; F - Frequent 26 - 50% cover; O - Occasional 11 - 25% cover; R - Rare 1 - 10% cover.

QUADRAT 1

Species Richness:
(How many plant types?)

Sward Height (cm):
(Maximum height of the tallest Plant in your quadrat)

What Species can you see (list below)?:	Frequency (D, A, F, O, R):

QUADRAT 2

Species Richness:
(How many plant types?)

Sward Height (cm):
(Maximum height of the tallest plant in your quadrat)

What Species can you see (list below)?:	Frequency (D, A, F, O, R):

QUADRAT 3

Species Richness:
(How many plant types?)

Sward Height (cm):
(Maximum height of the tallest plant in your quadrat)

What Species can you see (list below)?:	Frequency (D, A, F, O, R):

5. Northern Brown Argus and Rock Rose Surveys: Your next task is to help us locate areas where the Northern Brown Argus may already be or could be found. We ask you to tell us if you have seen any common rock-rose either within your quadrats, or on the site/along your transect. If there is common rock-rose present, please conduct an egg search, by looking for white eggs on the upper side of common rock-rose leaves. Record what you see below:

Common Rock-rose

Northern Brown Argus and egg

OS Grid Ref/Lat Long of Rock-rose Location	Rock-rose abundance e.g. common & widespread, sparse and local, rare	Number of Northern Brown Argus Eggs	Number of Northern Brown Argus Adults

6. Butterfly Survey: Here is where you note down the butterfly species you observed when walking your transect:

Butterfly Species	Number of Individuals Observed	Sex, if possible

Thank you for your participation!

Appendix E-4. Volunteer recruitment poster call out.

Ecosystem Explorers: Call for Volunteers!

Discover the hidden world of insects and their homes in all the green spaces around you! The UK's species-rich grasslands (those home to a variety of plant species) have been reduced by more than 97% over the past century, threatening the vital services that these habitats provide us (from carbon storage to pollination). The loss of these grasslands threatens associated wildlife, including many butterfly and moth species, and we must monitor and manage these vulnerable habitats and species to avoid their permanent loss.
Ecosystem Explorers is a novel citizen science survey that aims to combine satellite imagery with environmental observations to save our species-rich grasslands and its associated butterflies and moths…***but we need your help!***

We are asking participants to get exploring over the summer season and conduct butterfly and habitat transects to identify unrecorded areas of species-rich grasslands across Scotland. We are also interested in finding the locations of key species: in particular, common rock-rose for the conservation of the vulnerable Northern Brown Argus. During your survey you will:

1. Walk a transect and record the general grassland habitat and any observed butterflies.
2. Conduct three small quadrat surveys to record the floral diversity across the habitat.
3. Look for common rock-rose and Northern Brown Argus eggs on your site.

If you would like to learn more about the survey and how to get involved, please email the lead researcher Samantha Suter at   You can also follow the project's updates and learn how to get involved on the survey web platform:

https://citsci.org/projects/ecosystem-explorers.

To join the project, you will need to sign up to the platform (with a pseudonym if you would like to stay anonymised). More information will be provided on the online platform in the coming weeks with detailed video tutorials and instructions for participating!
Thank you and we look forward to having you involved and contributing to the conservation of the UK's butterflies and moths!

Appendix E-5. Example of outreach activities designed for increased public engagement with the citizen science survey, Ecosystem Explorers.

University of Glasgow

Glasgow Science Festival

Ecosystem Explorers

An ecosystem is the interaction between animals and plants and their habitat: the place in which they live. For wildlife to survive certain environmental conditions need to be met e.g. temperature, rainfall, food, shelter and more!

Unfortunately, many ecosystems are disappearing due to habitat loss causing many animals to lose their homes.

It is vital to protect different habitats to support the wildlife that relies on them. For example, insects that live in species-rich grasslands are needed for pollination, improving soil health, prey for other animals, and to indicate changes in the environment.

By learning more about the ecosystems around us, we can discover the array of wildlife communities and what they need to thrive on the Earth. We invite you to explore the ecosystems around you!

Instructions

Exploration can take as long as you like.

- Your first task is to identify different ecosystems that you come across around your local area.
- When you discover different ecosystems have a look around for the different microhabitats as well as the invertebrates you see in the area.
- Finally, have a think about the differences in the ecosystems and why different wildlife may choose to live there.

Kit List:

- This survey booklet
- ID guide (see next page and more available online)
- General knowledge or map of your local area
- A mobile phone/camera (optional)
- If using a phone: an app such as iNaturalist, SmartScan, or Seek to help you identify plants and insects

Ecosystems to look out for:

- Lawn
- Flowerbeds
- Hedges
- Woodland
- Ponds

Ecosystem 1: ....

Microhabitats (tick what you see):

- ☐ Soil
- ☐ Fallen Leaves
- ☐ Flowers
- ☐ Manmade surfaces

Invertebrates (list what you see):

____

____

____

____

Ecosystem 2: ....

Microhabitats (tick what you see):

- ☐ Soil
- ☐ Fallen Leaves
- ☐ Flowers
- ☐ Manmade surfaces

Invertebrates (list what you see):

____

____

____

____

Ecosystem 3: ....

Microhabitats (tick what you see):

- ☐ Soil
- ☐ Fallen Leaves
- ☐ Flowers
- ☐ Manmade surfaces

Invertebrates (list what you see):

____

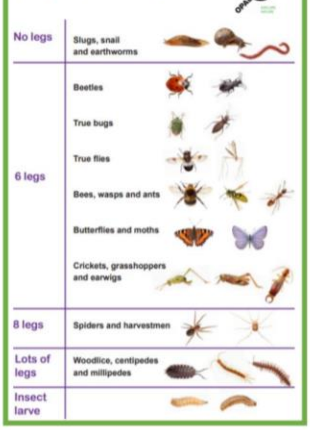
____

____

____

ID Guide

Use the guide below designed by OPAL to identify the creatures you find in each habitat.



Appendix E-6. Instructions for participating in Ecosystem Explorers, found on the online platform.

Ecosystem Explorers Guidelines

Hello and welcome to Ecosystem Explorers – a citizen science survey designed for locating species-rich grasslands across Scotland.

Here, we are using imagery from satellites to try locating species-rich grasslands…**but we need your help!**

We are asking you to tell us whether our satellite data has accurately mapped different classes of species-rich grasslands as well as help us confirm specific indicators for species-rich grassland identification. This is so we can conserve these vital habitats for priority species protection, such as the vulnerable Northern Brown Argus Butterfly.

Please follow the instructions below and watch the Instructions Video to help…

Getting Started:

1) Firstly, find the '<u>Grassland Survey Locations'</u> document. This is where you will find the maps that have been generated from the habitat classification model. The maps demonstrate the areas of grassland classes that have been determined from our model. However, we need to know if these maps are correct!
2) Using the map outputs, decide on a location you wish to survey. The map outputs will have sections of the Ordinance Survey that corresponds to the map output, as a georeference. There are also locations identified within the maps to help you locate the sites.
3) If you don't want to use the maps, there are a list of survey locations under the "data" tab > "locations". Here, you can find a list of coordinate locations across the country which can be used to guide your surveying locations!
4) Once a location has been decided, go out and survey! Using the other documents online print off hard copies of the survey form <u>('Grassland Classification Field Survey Form'</u>) to fill this out whilst you're on site. If you'd rather, you can download the CitSci.org app and follow the Ecosystem Explorers page to enter data online whilst you're in field!
5) Use the '<u>Grassland Classification Key Indicator' and 'ID guides and Resources'</u> documents to help you whilst you're surveying. You can familiarise yourself beforehand or you can print these off as a reference.

Conducting the Surveys:

1) When arrived on site, first fill in the location (OS grid reference and if possible, latitude and longitude from your phone), the weather conditions, and any first impressions of the site on the '<u>Grassland Classification Field Survey Form'</u>.
2) Select a suitable random location for your transect to start. Walk a 1 km transect, noting the grassland classes you see (using the '<u>Grassland Classification Key Indicators'</u> for help).
3) Whilst conducting the first walk, record any butterfly species within 5 m of either side of the transect. Record these on the survey form.
4) Once you have walked the first 1 km length of your transect, reverse this to conduct your quadrat surveys, taking a survey at the beginning (0 m), middle (500 m), and end (1 km) of your transect.
5) Place you're quadrat (use the length of an A4 piece of paper) at the associated section of your transect.
6) Within each quadrat record 4 randomly selected grass heights, measuring from the ground to the tip of the highest vegetation touching the ruler. Measure in cms.
7) Still in the quadrat, count the number of species you can see (this is known as species richness).
8) Finally, list any species that you can accurately identify in each quadrat. Common names are acceptable, and lower lever classifications e.g. oat grass vs false oat grass.
9) Conduct Egg surveys using the '<u>Northern Brown Argus Survey Guidance'</u> document to help.
10) Take any pictures of sites for reference e.g. wider habitat, specific plant and butterfly species, eggs seen.

Entering the data:

1) Once you have filled in the survey forms, this data must be entered online if you did not do it on the app.
2) Go to the data entry forms tab.
3) Enter data online from your associated survey sheet e.g. Grassland Data entry Form online = Grassland Classification Field Survey Form.

Appendix E-7. Excerpt of Ecosystem Explorers survey event on Eventbrite (2023) and associated survey dates.

Sunday, 9 July

Ecosystem Explorers: Hill of Tillymorgan

Discover the hidden world of insects and their homes in all the green spaces around you.

About this event

Ecosystem explorers invites nature lovers to join us in searching for species-rich grasslands, as part of our Citizen Science survey.

Take part in our botanical and butterfly surveys in Scotland's countryside to protect a threatened habitat and associated vulnerable wildlife.

Spend the day outside, improve your biodiversity knowledge, and meet like-minded people!

Resources and more information can be found by signing up to the project here: https://www.citsci.org/projects/ecosystem-explorers

We will:

1) Walk a 1 km transect to determine the grassland type

2) Conduct a butterfly survey along the transect

3) Conduct 3 small quadrat surveys to identify specific wildflowers

4) Search for priority species: Common Rock-rose and Northern Brown Argus butterfly eggs

To note:

Parking - 57.393016, -2.584680 Parking just past, use coords in google maps, Kirkton of Culsalmond Old Parish Church, Lawrence Rd, Insch AB52. Please car share where possible. Any issues call Sam on

Terrain - Possible uneven or rough footpaths, wear appropriate footwear (walking boots) for any walking through fields.

Length - We may be walking approximately 7 km. Slight gradual ascent <200 m.

(NB. We are working with prediction maps and finding species-rich grasslands to survey is not guaranteed. In the event of this, the day will still be fun and educational, a chance to meet other people, and explore Scotland's countryside).

Frequently asked questions

How long is the day?

Date	Location
13th June 2023	Menstrie Glen
22nd June 2023	Fallin Bing
8th July 2023	Green Knowes Wind Farm
9th July 2023	Hill of Tillymorgan
17th July 2023	Kincraig, Cairngorms
18th July 2023	Dunnet Head
20th July 2023	Keoldale headland
4th August 2023	Campsie Fells

Appendix E-8 - R code

##Chi-squared goodness of fit tests

```{r}

total <- c(X,Y) ##where X = number of model predicted SRG pixels and Y = number of corresponding citizen observed SRG locations

model <- chisq.test(total, p = c(1/2, 1/2)) ## 1/2 = the theoretical split through the assumption that the model predictions and the citizen observations should align for each specific site/pixel

summary(model)

```

##Chi squared test for independence/fisher's exact tests

(Using the example of the assessing the agreement level between the model predictions and the citizen observations by participant group, the rows are defined by the grouping factor whilst the columns are defined as the numbers across the agreement levels (X, Y, Z))

```{r}

data <- data.frame(

  “Ecosystem Explorers” = c(X,Y,Z),

  « Nature Scot » = c(X,Y,Z),

  « Plantlife » = c(X,Y,Z),

  « BSBI » = c(X,Y,Z),


  row.names = c(“no match”, “partial match”, “full match”),

  stringsAsFactors = FALSE

)

colnames(data) <- c(“Ecosystem Explorers”, “Nature Scot”, “Plantlife”, “BSBI”)

data

test <- fisher.test(data)

summary(test)

```

Appendix F

Appendix F-1. Documentation for project implementation into schools and how it reaches curriculum targets in Biography, Geography, and IT.

Project Overview

Rising human population has led to increased destructive activities from higher demands for resources, such as food. These activities have instigated a global environmental crisis. For example, the UK underwent large agricultural industrialisation in the later half of the 20th century. This intensification led to the loss of several natural and semi-natural habitats, so much so that now over 70% of the UK land cover is made up of agricultural land. As a result, this has caused widespread native species decline in much of the UK's flora and fauna. A habitat found in the UK that has been reduced by more than 97% is that of species-rich grasslands, which now make up approximately 1% of the UK land cover. This is a semi-natural habitat with a high floral species diversity which provides ecosystems services such as carbon sequestration, as well as vital habitat for many priority species. One such species is the Northern Brown Argus butterfly, which has seen population declines of nearly 60% and is now protected under UK legislation. To try halt and reverse any damaging impacts on future habitats and species in the UK, habitat extent and species population trends must be monitored to ensure accurate management and conservation. However, species-rich grasslands in the UK are not mapped extensively, making it difficult to target conservation efforts to these habitats and their species. As well as this, ensuring the correct monitoring and management is time and labour intensive. Therefore, a tool must be designed to enable sufficient monitoring of this habitat to ensure its protection and help reduce the negative impacts on the Northern Brown Argus. This project aims to bring citizens together to identify areas of species-rich grassland from satellite imagery, so conservation can be targeted to these locations and help the Northern Brown Argus and other priority species. Hopefully, it will also provide a framework for an approach that can be directed towards other declining habitats and species. To enable this, the project will be working openly, meaning that the entire process will be available to other researchers, the public, organisations, and governments. This is done by using OS practices, databases, and software.

Project in Curriculum

This project can be instigated into the national curriculum for Scotland as its interdisciplinary approach can target multiple subject areas (such as biology, geography/environmental science, technology, and citizenship). The project would raise awareness about a current biodiversity issue in Scotland because of human action, the importance of conservation and how to close gaps in knowledge, how to use various methods and technologies to tackle a topical issue, and how to work openly and reproducibly to increase validity and trustworthiness in research.

S3 Curriculum Targets

Project Process

1) *The habitat of concern must be classified and sites suitable for in situ field measurements to support satellite imagery must be located. This is to increase identification accuracy. These initial sites must be where the Northern Brown Argus is found, and as such the appropriate habitat conditions must be met. These include having the presence of their larval food plant, the common-rock rose on south-facing slopes of altitudes less than 350 m. Precise climatic conditions are unknown. Locating these sites is conducted using ArcGIS pro. Open access data can be found for the distribution of Northern Brown Argus, Common Rock Rose, elevation data, and habitat classes found in Scotland. Overlaying these features will narrow results for site selection. It is also important to find areas where the Northern Brown Argus is not found but its habitat requirements appear to be suitable. It can then be investigated in the field as to why the butterfly is not present.*

In this section of the project, multiple areas of the curriculum can be targeted. Specifically, Social Sciences and Technologies would be addressed largely here. In Social Sciences, codes **SOC 3-10a**, and **SOC 3-14a** will be reached by using GIS to "investigate climate, physical features [such as elevation], and living things [rock rose and Northern Brown Argus] of a natural environment [species-rich grassland] different from [the students'] own and explain their interrelationship" by searching for sites where these features intersect. This will "use a range of maps and geographic information systems to gather, interpret, and present conclusions". It will combine with Technologies codes **TCH 3-01a** and **TCH 3-02a**, as the use of software such as GIS and open data platforms (such as NBA Atlas) will "use the features of a range of digital technologies, integrated software and online resources to determine the most appropriate to solve problems" as the students will use "digital technologies to search, access and retrieve information”. Retrieving data from open data repositories will also give students an “awareness of plagiarism”.

2) *Satellite imagery must be obtained. Certain satellites are now under open licenses and accessing imagery is free, such as Landsat and Sentinel data. These images must be pre-processed and combined with in situ field measurements. Coding programmes that are also freely available such as R or Python can be used to process these images. A library of images will be collated to use for identification of species-rich grasslands.*

This will target the codes highlighted above as it further uses various software, resources, and geographic information systems to investigate the environmental features by another technique.

3) *A citizen science platform will be developed on a website such as citsci.org. Here, participants can access information about the purpose of the project, how they can help, how to identify species-rich grasslands from satellite imagery that is supplied on the platforms, and further information such as results, community forums, and in person events.*

This will address the code in Social Science **SOC3-08a**, as students will help “identify possible consequences of an environmental issue [species-rich grassland loss] and make informed suggestions about ways to manage the impact” from locating sites of conservation importance. Technology codes **TCH 3-07a** and **TCH 3-15a** are met through using various software and websites students will be able to evaluate the “costs and benefits of using technologies to reduce the impact of our activities on the environment” and “select appropriate development tools to design, build, evaluate and refine computing solutions based on requirements” of an issue by experiencing the variety of online resources.

4) *After the imagery is analysed by participants, the accuracy of habitat identifications must be assessed. This is done through ground truthing where surveyors will go to identified areas of species-rich grasslands and confirm whether or not the identification was correct through standard surveying techniques such as phase 1 habitats.*

Through the process of the project, students should learn about various biological systems which can target the code in Sciences **SCN 3-01a**. Students could “identify living things from different habitats to compare their biodiversity and can suggest reasons for their distribution” from information that is available on the citizen science project platform, where biological resources and keys will be available related to the project. Ground truthing will also include surveying and sampling techniques of biodiversity in the located habitats to reach this outcome in the curriculum.

5) *The information from the project will be disseminated via various social media platforms to raise awareness and increase participation. Outputs from the research process will be openly available on appropriate platforms.*

Having an awareness of the different OS systems will tackle specifically the technology codes mentioned above to evaluate various digital sources for a variety of purposes.

S4 Curriculum Targets

Project Process

6) *The habitat of concern must be classified and sites suitable for in situ field measurements to support satellite imagery must be located. This is to increase identification accuracy. These initial sites must be where the Northern Brown Argus is found, and as such the appropriate habitat conditions must be met. These include having the presence of their larval food plant, the common-rock rose on south-facing slopes of altitudes less than 350 m. Precise climatic conditions are unknown. Locating these sites is conducted using ArcGIS pro. Open access data can be found for the distribution of Northern Brown Argus, Common Rock Rose, elevation data, and habitat classes found in Scotland. Overlaying these features will narrow results for site selection. It is also important to find areas where the Northern Brown Argus is not found but its habitat requirements appear to be suitable. It can then be investigated in the field as to why the butterfly is not present.*

In this section of the project, multiple areas of the curriculum can be targeted. Specifically, Social Sciences and Technologies would be addressed largely here. In Social Sciences, codes **SOC 4-14a** will be reached by using “geographical information systems [ArcGIS] to identify patterns of human activity and physical processes”. It will combine with Technologies codes **TCH 4-01a** and **TCH 4-02a**, as the use of software such as GIS and open data platforms (such as NBA Atlas) will “use digital technologies to access, select relevant information and solve real world problems” as the students will use “digital technologies to process and manage information responsibly and can reference sources accordingly2. Retrieving data from open data repositories will also give students insight into how to “reference sources accordingly”.

7) *Satellite imagery must be obtained. Certain satellites are now under open licenses and accessing imagery is free, such as Landsat and Sentinel data. These images must be pre-processed and combined with in situ field measurements. Coding programmes that are also freely available such as R or Python can be used to process these images. A library of images will be collated to use for identification of species-rich grasslands.*

This will target the codes highlighted above as it further uses various software, resources, and geographic information systems to investigate the environmental features by another technique.

8) *A citizen science platform will be developed on a website such as citsci.org. Here, participants can access information about the purpose of the project, how they can help, how to identify species-rich grasslands from satellite imagery that is supplied on the platforms, and further information such as results, community forums, and in person events.*

This will address the code in Social Science **SOC4-08a and SOC 4-10a**, as students will help address “the sustainability of key natural resources and analyse the possible implications for human activity” from locating sites of conservation importance and associated ecosystem services. Students will “develop [their] understanding of the interaction between humans and the environment by describing and assessing the impact of human activity on an area” through learning about the importance of this project in participation. Technology code **TCH 4-15a** are met through using various software and websites students will be able to evaluate the “costs and benefits of using technologies to reduce the impact of our activities on the environment” and “select appropriate development tools to design, build, evaluate and refine computing solutions to process and present information whilst making reasoned arguments to justify my decisions” of an issue by experiencing the variety of online resources.

9) *After the imagery is analysed by participants, the accuracy of habitat identifications must be assessed. This is done through ground truthing where surveyors will go to identified areas of species-rich grasslands and confirm whether or not the identification was correct through standard surveying techniques such as phase 1 habitats.*

Through the process of the project, students should learn about various biological systems which can target the code in Sciences **SCN 4-01a**. Students should “understand how animal and plant species depend on each other and how living things are adapted for survival.” from information that is available on the citizen science project platform and seeing how certain factors (such as larval food presence) influences other factors (such as butterfly occurrence), where biological resources and keys will be available related to the project.

10) *The information from the project will be disseminated via various social media platforms to raise awareness and increase participation. Outputs from the research process will be openly available on appropriate platforms.*

Having an awareness of the different OS systems will tackle specifically the technology codes mentioned above to evaluate various digital sources for a variety of purposes.

Highers Curriculum Targets

Project Process

11) *The habitat of concern must be classified and sites suitable for in situ field measurements to support satellite imagery must be located. This is to increase identification accuracy. These initial sites must be where the Northern Brown Argus is found, and as such the appropriate habitat conditions must be met. These include having the presence of their larval food plant, the common-rock rose on south-facing slopes of altitudes less than 350 m. Precise climatic conditions are unknown. Locating these sites is conducted using ArcGIS pro. Open access data can be found for the distribution of Northern Brown Argus, Common Rock Rose, elevation data, and habitat classes found in Scotland. Overlaying these features will narrow results for site selection. It is also important to find areas where the Northern Brown Argus is not found but its habitat requirements appear to be suitable. It can then be investigated in the field as to why the butterfly is not present.*
12) *Satellite imagery must be obtained. Certain satellites are now under open licenses and accessing imagery is free, such as Landsat and Sentinel data. These images must be pre-processed and combined with in situ field measurements. Coding programmes that are also freely available such as R or Python can be used to process these images. A library of images will be collated to use for identification of species-rich grasslands.*

13) *A citizen science platform will be developed on a website such as citsci.org. Here, participants can access information about the purpose of the project, how they can help, how to identify species-rich grasslands from satellite imagery that is supplied on the platforms, and further information such as results, community forums, and in person events.*
14) *After the imagery is analysed by participants, the accuracy of habitat identifications must be assessed. This is done through ground truthing where surveyors will go to identified areas of species-rich grasslands and confirm whether or not the identification was correct through standard surveying techniques such as phase 1 habitats.*
15) *The information from the project will be disseminated via various social media platforms to raise awareness and increase participation. Outputs from the research process will be openly available on appropriate platforms.*

Meeting the Curriculum

Areas of Higher Geography, Biology, and potentially Computing Science can be met through the various project steps outlined above. In this section of the project, knowledge on areas of the course specification for both Higher Biology and Geography can be met. For Higher Biology "components of biodiversity" and "threats to biodiversity" will be specifically targeted. Mapping species-rich grasslands through remote sensing and identifying sites through ArcGIS will show how the habitat has been reduced to small fragmentations to show how "habitat loss…impact[s] on species richness". Identifying sites where the Northern Brown Argus populations have decreased, increased, and are stable will demonstrate how species richness varies so that fragmentation "may result in a decrease in biodiversity" (**Higher Biology**). Information on identification keys will further knowledge on species richness and ground truthing will demonstrate variance in this and diversity by assessing habitats *in situ* after identification through the citizen science survey. Higher Geography will be targeted through "global issues…to demonstrate the interaction of physical and human factors". Obtaining data from various sources (data repositories) and interpreting and analysing data in multiple software e.g. ArcGIS/ the citizen science platform) and using these to identify site locations in ArcGIS will be relevant for "researching and evaluating a wide range of information collected from a range of sources about complex geographical issues" as well as for "using a wide range of mapping skills and techniques in geographical contexts which may be familiar or unfamiliar". This will target "information handling" as well as "citizenship" aspects (**Higher Geography**). Higher Computing Science may be targeted through web design.

Suggested Projects/Activities

Higher Biology – develop their own citizen science project around a protected species/habitat/assemblage. Create species richness/diversity surveys.
Higher Geography – mapping projects. Target another declining habitat through remote sensing/GIS.

Advanced Highers Curriculum Targets

Project Process

16) *The habitat of concern must be classified and sites suitable for in situ field measurements to support satellite imagery must be located. This is to increase identification accuracy. These initial sites must be where the Northern Brown Argus is found, and as such the appropriate habitat conditions must be met. These include having the presence of their larval food plant, the common-rock rose on south-facing slopes of altitudes less than 350 m. Precise climatic conditions are unknown. Locating these sites is conducted using ArcGIS pro. Open access data can be found for the distribution of Northern Brown Argus, Common Rock Rose, elevation data, and habitat classes found in Scotland. Overlaying these features will narrow results for site selection. It is also important to find areas where the Northern Brown Argus is not found but its habitat requirements appear to be suitable. It can then be investigated in the field as to why the butterfly is not present.*
17) *Satellite imagery must be obtained. Certain satellites are now under open licenses and accessing imagery is free, such as Landsat and Sentinel data. These images must be pre-processed and combined with in situ field measurements. Coding programmes that are also freely available such as R or Python can be used to process these images. A library of images will be collated to use for identification of species-rich grasslands.*
18) *A citizen science platform will be developed on a website such as citsci.org. Here, participants can access information about the purpose of the project, how they can help, how to identify species-rich grasslands from satellite imagery that is supplied on the platforms, and further information such as results, community forums, and in person events.*
19) *After the imagery is analysed by participants, the accuracy of habitat identifications*

must be assessed. This is done through ground truthing where surveyors will go to identified areas of species-rich grasslands and confirm whether or not the identification was correct through standard surveying techniques such as phase 1 habitats.

20) *The information from the project will be disseminated via various social media platforms to raise awareness and increase participation. Outputs from the research process will be openly available on appropriate platforms.*

Meeting the Curriculum

Areas of Higher Geography, Biology, and potentially Computing Science can be met through the various project steps outlined above. In this section of the project, knowledge on areas of the course specification for both Advanced Higher Biology and Geography can be met. For Advanced Higher Biology "organisms and evolution" will be specifically targeted. This will be through the variety of sampling techniques that are incorporated into the project. When taking *in situ* field measurements at the start of the project and when completing ground truthing after the citizen science survey, this will include information on "sampling wild organisms" and "vulnerable species and habitats, which are protected by legislation". Identification keys will be available as part of the citizen science platform to target "Identification and taxonomy Identification of a sample…using classification guides, biological keys". As Northern Brown Argus is an indicator species of species-rich grasslands "monitoring populations, presence, absence or abundance of indicator species can give information of environmental qualities" in this case site condition. The platform of the citizen science project and advertisement of the project and dissemination of results will target "Scientific literature and communication [to explain] the importance of publication of methods, data, analysis and conclusions in scientific reports", whilst the open nature of the project will promote various practices in the research process. Advanced Higher Geography will be targeted through the various mapping systems, obtaining data from multiple sources (data repositories) and interpreting and analysing data in software e.g. ArcGIS/ the citizen science platform to "demonstrate mapping skills techniques through their ability to use evidence from maps and other supplementary items". The citizen science survey will show students methods to "interpret and use information from supplementary items such as maps or map-based diagrams". Certain parts of the project such as *in situ* measurements will demonstrate various techniques such as "Vegetation sampling…to determine the amount and variety of vegetation in a prescribed area, for example: vegetation amount — a quadrat, randomly thrown, and number of species per square". Advanced Higher Computing Science may be targeted through web design/coding?

Suggested Projects/Activities

Advanced Higher Biology – develop their own citizen science project around a protected species/habitat/assemblage. Create species richness/diversity surveys.

Advanced Higher Geography – mapping projects. Target another declining habitat through remote sensing/GIS. In situ habitat measurements, vegetation sampling techniques, soil, slope analysis.