Mapping Environmental Factors to *Fasciola hepatica* at Farm-Level

Qi Qin

Submitted in fulfilment of the requirements for
the degree of Master of Science in Earth Science

School of Geographical and Earth Science
College of Science and Engineering
University of Glasgow

November 2018

Abstract

Fasciola hepatica is a common parasite, which affects sheep and cattle. In past decades, due to wet weather and the increasing temperature in summer and early autumn, the risk of sheep and cattle infection by *F. hepatica* has significantly increased, and it has also caused enormous economic loss. The distribution of *F. hepatica* is profoundly affected by environmental conditions, including precipitation, temperature, soil moisture, and vegetation type and can be predicted on a regional scale by environmental data. In the UK, a large number of farms are small or medium-sized, and the current parasite forecast system works on a national scale postal area scales, which is too broad to provide a precise forecast or enable the development of within-farm management strategies. Hence, there is a need to map the environment data to *F. hepatica* at the farm-level.

This study examined the relationships between *F. hepatica* and environmental factors at the within-farm and within-field levels, useing earth observation techniques and geographic information systems (GIS) to analyse the potential contribution of environmental factors to the density of *F. hepatica* on pasture. Five study fields within two farms in Scotland, Cochno and Dumgoyne, were used in this study. Nine variables grouped within four categories were used to describe the environmental conditions: 1) remote sensing indices: normalized difference water index (NDWI) and normalized difference vegetation index (NDVI); 2) soil properties: soil moisture and soil temperature; 3) topographical factors: elevation, slope, and aspect; 4) grazing grass: grass height and grass weight. Two variables were used to describe the intensity of *F. hepatica* in pasture grass: count of metacercarial cysts and yield of metacercariae (the number of metacercarial cysts per gram pasture). Univariate negative binomial regression models of each environment factors were built against the number of metacercariae per 0.1 gram. Subsequently, the significant variables were used in a multivariate regression model. The environmental conditions were diverse among the five fields and the *F. hepatica* density also varied quite widely. The results of this study showed that there were strong associations among environmental variables. These associations result in complex interactions when modelling the environment to *F. hepatica*. Relationships among environmental variables and metacercarial density measures were quite inconsistent. Although the NDVI, grass weight, soil moisture and slope in particular fields showed significant relationships with *F. hepatica*, the remote sensing indicators tested would not be useful for determining parasite concentrations within farm-level.

Acknowledgements

I would first like to thank my supervisors Dr. Brian Barrett and Prof. Nicholas Jonsson for their help and support on this project. They consistently allowed this paper to be my own work but steered me in the right direction whenever they thought I needed it. I would also like to thank Mr. James McGoldrick who supported my work in the laboratory to process the grass samples.

Declaration

I, Qi Qin, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Qi Qin

12 November 2018

Contents

Abstract i

Acknowledgements iii

Declaration v

1 Introduction 1

1.1 Introduction . . . 1

1.2 *F. hepatica* and fasciolosis . . . 2

1.2.1 Biological characteristics of *F. hepatica* . . . 2

1.2.2 Environmental effects on the life-cycle of *F. hepatica* . . . 6

1.2.2.1 Climate . . . 6

1.2.2.2 Soil moisture and soil material . . . 7

1.2.2.3 Landcover type and other topographical factors . . . 7

1.3 Remote sensing and GIS approaches used in Parasitology . . . 8

1.4 Aims and objectives . . . 9

2 Methodology 11

2.1 Study Area . . . 11

2.2 Earth Observation Imagery . . . 13

2.2.1 Sentinel-2 . . . 13

2.2.1.1 Pre-processing of satellite images . . . 14

2.2.2 LiDAR . . . 15

2.2.3 Aerial photography . . . 16

2.3 Reference data . . . 17

2.3.1 Terrain data . . . 17

2.3.2 Soil texture . . . 22

2.4 Remote sensing indices . . . 23

2.4.1 Normalized difference water index . . . 23

2.4.2 Normalized difference vegetation index . . . 24

2.5 Preliminary risk map . . . 24

2.6 Field survey and laboratory analysis . . . 25
2.6.1 Field survey method . . . 25
2.6.1.1 Selection of sampling points . . . 25
2.6.1.2 Materials . . . 25
2.6.1.3 Sampling approach . . . 26
2.6.2 Laboratory analysis . . . 28
2.6.2.1 Materials . . . 28
2.6.2.2 Working Procedures . . . 29
2.6.3 Calculation of metacercarial cyst count . . . 30
2.7 Statistical analysis methods . . . 31
2.7.1 Preliminary analysis . . . 31
2.7.2 Kruskal-Wallis one-way analysis of variance . . . 32
2.7.3 One-way analysis of variance (ANOVA) . . . 33
2.7.4 Pearson's Correlation Coefficients . . . 33
2.7.5 Modelling methods to assess relationships between *F. hepatica* and environmental factors . . . 34
2.7.5.1 Negative binomial regression model . . . 34
2.7.5.2 McFadden's Pseudo R^2 . . . 35
2.7.5.3 Akaike information criterion (AIC) . . . 36
2.7.5.4 Chi-Squared goodness of fit test . . . 36

3 Preliminary Risk Maps 39
3.1 Risk classification . . . 39
3.2 Sampling locations . . . 43

4 Exploratory Analysis of Environmental Factors and Density of *F. hepatica* in pasture 47
4.1 Initial exploration of *F. hepatica* and environmental factors among fields . . . 47
4.1.1 *F. hepatica* density in pasture . . . 47
4.1.1.1 Metacercarial cyst counts . . . 47
4.1.1.2 Yield (Metacercariae/g) . . . 52
4.1.2 Remote sensing indices . . . 57
4.1.2.1 Normalized difference water index . . . 57
4.1.2.2 Normalized difference vegetation index . . . 62
4.1.3 Topographical factors . . . 67
4.1.3.1 Elevation . . . 67
4.1.3.2 Slope . . . 72
4.1.3.3 Aspect . . . 77
4.1.4 Soil properties . . . 81

4.1.4.1 Soil Moisture . . . 81
4.1.4.2 Soil Temperature . . . 86
4.1.5 Pasture . . . 91
4.1.5.1 Grass Height . . . 91
4.1.5.2 Grass Weight . . . 96
4.2 Relationships between remote sensing, terrain, soil and pasture datasets . . . 101
4.2.1 Cochno Low . . . 101
4.2.2 Cochno Mid . . . 102
4.2.3 Cochno High . . . 102
4.2.4 Dumgoyne North . . . 103
4.2.5 Dumgoyne South . . . 104

5 Relationships between *F. hepatica* and Environmental Factors 105
5.1 Univariate regression models . . . 105
5.1.1 Modeling remote sensing indices against *F. hepatica* density in pasture 105
5.1.2 Modeling Soil properties against *F. hepatica* density in pasture . . . 108
5.1.3 Modeling topographical factors against *F. hepatica* density in pasture . 111
5.1.4 Modeling pasture grass against *F. hepatica* density in pasture . . . 115
5.2 Multivariate regression model . . . 118

6 Discussion and conclusion 121
6.1 Discussion . . . 121
6.2 Conclusion . . . 123

Bibliography 125

Appendix A Monthly precipitation and temperature of the study areas 133

Appendix B Sampling Locations of each field 135

Appendix C *F. hepatica* and environmental factors of each sampling point 137

Appendix D Summary of statistical results in Chapter 4 143

Appendix E Google Earth Engine JavaScript Codes 145

List of Tables

2.1 Characteristics of the fields at the two study locations 12
2.2 Spectral bands for the SENTINEL-2 sensors(S2A) used in the study 14

3.1 Risk class for the data layers (Risk decreasing from 5 to 1) 40
3.2 Soil texture of study areas and the risk class (Risk decreasing from 7 to 1) . . . 40

4.1 Pearson's Correlation Coefficient of Cochno Low 102
4.2 Pearson's Correlation Coefficient of Cochno Mid 102
4.3 Pearson's Correlation Coefficient of Cochno High 103
4.4 Pearson's Correlation Coefficient of Dumgoyne North 104
4.5 Pearson's Correlation Coefficient of Dumgoyne South 104

5.1 Univariate regression model of NDWI 106
5.2 Univariate regression model of NDVI 106
5.3 Chi-squared goodness of fit for remote sensing index models 108
5.4 Univariate regression model of Soil Moisture 109
5.5 Univariate regression model of Soil Temperature 109
5.6 Chi-squared goodness of fit for soil models 111
5.7 Univariate regression model of Elevation 112
5.8 Univariate regression model of Slope 112
5.9 Univariate regression model of Aspect 112
5.10 Chi-squared goodness of fit for topographical factors models 114
5.11 Univariate regression model of Grass Height 115
5.12 Univariate regression model of Grass Weight 116
5.13 Chi-squared goodness of fit for forage models 117
5.14 Multivariate regression model 118
5.15 Chi-squared goodness of fit for multivariate regression model 119

B.1 Sampling locations of Cochno 135
B.2 Sampling locations of Dumgoyne 136

C.1 CochnoLow data 137

C.2 CochnoMid data 138
C.3 CochnoHigh Data 139
C.4 DumgoyneNorth Data 140
C.5 DumgoyneSouth Data 141

D.1 Shapiro-Wilk normality test results in Chapter4 143
D.2 Kruskal-Wallis test results in Chapter4 143
D.3 One-way ANOVA test results in Chapter4 144

List of Figures

1.1 The life-cycle of *F.hepatica* (Taylor et al., 2016) 3
1.2 Metacercariae of *F. hepatica* (Photo by Mr.James McGoldrick) 4
1.3 Adult *F. hepatica* (up to 30 mm by 13 mm) (Photo by Mr.James McGoldrick) . 5

2.1 Maps of study areas: (a) Two study areas; (b) Five research fields. 12
2.2 Google Earth Engine API. 14
2.3 Workflow of pre-processing of satellite images. 15
2.4 LiDAR images: (a) Scotland 1 m LiDAR composites; (b) LiDAR tiles used in this study. 16
2.5 Aerial photography: (a) Cochno; (b) Dumgoyne. 17
2.6 Workflow of pre-processing of LiDAR DTM Data. 18
2.7 Elevation maps of research areas: (a) Cochno; (b) Dumgoyne. 19
2.8 Slope maps of research areas: (a) Cochno;(b) Dumgoyne. 21
2.9 Aspect maps of research areas: (a) Cochno; (b) Dumgoyne. 22
2.10 Workflow of field survey and quadrat sampling. 26
2.11 Field survey 28
2.12 Workflow of laboratory analysis. 30

3.1 Risk Classification Maps of Cochno: (a) NDWI; (b) NDVI; (c) Slope; (d) Soil. 41
3.2 Risk Classification Maps of Dumgoyne: (a) NDWI; (b) NDVI; (c) Slope; (d) Soil. 42
3.3 Risk Maps: (a) Cochno; (b) Dumyoyne. 43
3.4 Risk maps and metacercarial cyst counts: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South. 44

4.1 The metacercarial cyst count maps for sampling points within the study fields: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South. 48
4.2 Histograms of metacercarial cyst for each field and all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples. 51

4.3 Density curve of metacercarial cysts for each field. Red vertical line is the overall average count. Blue vertical line is the median. 52
4.4 The yield for sampling points within the study fields: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South. . . 53
4.5 Histograms of yield for each field and for all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples. 56
4.6 Density curve of yield (metacercariae/g) for each field. 56
4.7 The NDWI maps for study fields and the NDWI of each sampling point: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South. 58
4.8 Histograms of NDWI for each field and all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples. 61
4.9 Density curve of NDWI for each field. 61
4.10 The NDVI maps for the study fields and the NDWI of each sampling point: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South. 63
4.11 Histograms of NDVI for each field and for all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples. 66
4.12 Density curve of NDVI for each field. 66
4.13 The elevation maps for the study fields and the altitude of each sampling point: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South. 68
4.14 Histograms of elevation for each field and all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples. 71
4.15 Density curve of elevation for each field. 71
4.16 The slope maps for the study fields and the slope of each sampling point: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South. 73
4.17 Histograms of slope for each field and all sample: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples. 76
4.18 Density curve of slope for each field. 76
4.19 Histogram of aspect for each field. 77

4.20 The aspect maps for the study fields and the aspect of each sampling point: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South. 78
4.21 The soil moisture values for sampling points within the study fields: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South. 82
4.22 Histograms of soil moisture for each field and all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples. 85
4.23 Density curve of soil moisture for each field. 85
4.24 The soil temperature (degrees Celsius) for sampling points within the study fields: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South. 87
4.25 Histograms of soil temperature (degrees Celsius) for each field and for all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples. 90
4.26 Density curve of soil temperature for each field. 90
4.27 The grass height (cm) for sampling points within the study fields: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South. 92
4.28 Histograms of grass height (cm) for each field and all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples. 95
4.29 Density curve of grass height for each field. 95
4.30 The grass weight values for sampling points within the study fields: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South. 97
4.31 Histograms of grass height for each field and all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples. 100
4.32 Density curve of grass weight for each field. 100

5.1 The dot plot and fitted line of NDWI against Yield2. 107
5.2 The dot plot and fitted line of NDVI against Yield2. 107
5.3 The dot plot and fitted line of Soil Moisture against Yield2. 110
5.4 The dot plot and fitted line of Soil Temperature against Yield2. 110
5.5 The dot plot and fitted line of Elevation against Yield2 for each field. 113
5.6 The dot plot and fitted line of Slope against Yield2 for each field. 113
5.7 The dot plot of Aspect against Yield2. 114

5.8 The dot plot and fitted line of Grass Height against Yield2. 116
5.9 The dot plot and fitted line of Grass Weight against Yield2. 117

A.1 Monthly precipitation (mm) and temperature (degrees Celsius) of the study areas 134

Chapter 1

Introduction

1.1 Introduction

Fasciola hepatica, also known as liver fluke, is a parasitic trematode (flatworm or fluke) which is distributed world-wide. It causes the disease fasciolosis, among mammals and ruminants, including humans. Fasciolosis is a widespread disease around the world and causes huge financial losses to farmers and consumers in the form of poor quality carcass, reduction in growth rate and lower productivity (Khan et al., 2013). Notable economic loses occur due to infection of animals with *Fasciola*, with worldwide losses in animal production estimated to be approximately 3.2 billion US dollars per year (Mehmood et al., 2017). The World Health Organisation (WHO) also identified fasciolosis as a re-emerged, neglected tropical disease associated with human disease epidemics (Beesley et al., 2017). In Europe, both the awareness and prevalence of fasciolosis have increased.The significant impact on agriculture and human health, and the growing demand for animal-based foods that support the growth of the global population, make fasciolosis a major health problem. The prevalence of *Fasciolia* has been a issue for a long time. Fasciolosis caused by *F. gigantica* is prevalent in African, Asian and some American countries, and prevalence is comparatively higher in developing countries compared with developed countries in every continent (Jaja et al., 2017; Hussein and Khalifa, 2010). It is mentioned that the high prevalence rate is due to poor management practices and lack of knowledge among the farmers about its control. In Europe, *F. hepatica* is mostly associated with disease in sheep, cattle and goats.

In the UK, cattle infection by *F. hepatica* is estimated to cost £23 million annually. At the farm level, it has been reported that fasciolosis has an unquantifiable effect on milk production, carcass composition and extending the time to reach slaughter weight (http://www.nadis.org.uk/). Over the past decade, the regional epidemiology of fasciolosis in the United Kingdom has changed in terms of seasonality and severity. These changes are related to increased rainfall or localised flooding, prompting debate on the harmful effects of climate change (Sucke, 2014).

As a consequence, effective management of fasciolosis has become problematic in areas where fluke exists, leading to serious loss of production in sheep and cattle. Meanwhile, In the eastern parts of the UK, unexpected disease outbreaks have caused production losses (Sargison and Scott, 2011). This has prompted the investigation of advanced technologies and strategies in farming management. Monitoring changes in climate and grassland in real time might help to identify the potentially high risk areas or times for *F. hepatica*, which might reduce the number of infected livestock. For example, the National Animal Disease Information Service (NADIS) publishes a parasite forecast report every month and suggests "safe grazing" before predicted high risk months such as August to October (NADIS, 2017). In Scotland, fluke has traditionally been seen in the wetter western areas, however, over recent years, cases have been reported in the eastern and north-eastern regions in increasing numbers (Mitchell, 2002; Kenyon et al., 2009).

1.2 *F. hepatica* and fasciolosis

1.2.1 Biological characteristics of *F. hepatica*

F. hepatica is an extraordinarily successful parasite that is found in most temperate regions of the world. It has an indirect life cycle involving a final host (in which the adult fluke reproduces), an intermediate host (in which the larval stages of the worm develop) and a carrier (entailing suitable aquatic plants). Adult stages of *F. hepatica* flukes occur in the liver of the final host, which include cattle, sheep, and buffaloes. Wild ruminants and other mammals, including humans, can act as definitive hosts as well. The molluscan intermediate hosts of *F. hepatica* are amphibious snails from the family *Lymnaeidae* and the principal species in Europe is *Galba truncatula* (Torgerson and Claxton, 1999).

The life cycle (Figure 1.1) starts when infected animals defecate in fresh-water sources. Undifferentiated fluke eggs are passed out in the feces of infected animals and once washed out of the feces, the eggs start to develop, a process dependent on temperature. When a fully developed egg is given stimuli of increased light and temperature, the miracidium is released. It requires water to swim through and once it finds a suitable lymnaeid snail, the miracidium burrows through the foot and into the snail body and develops into sporocyst larva (around 0.7 mm). Each sporocysts then ruptures its body wall and emerges as two rediae.

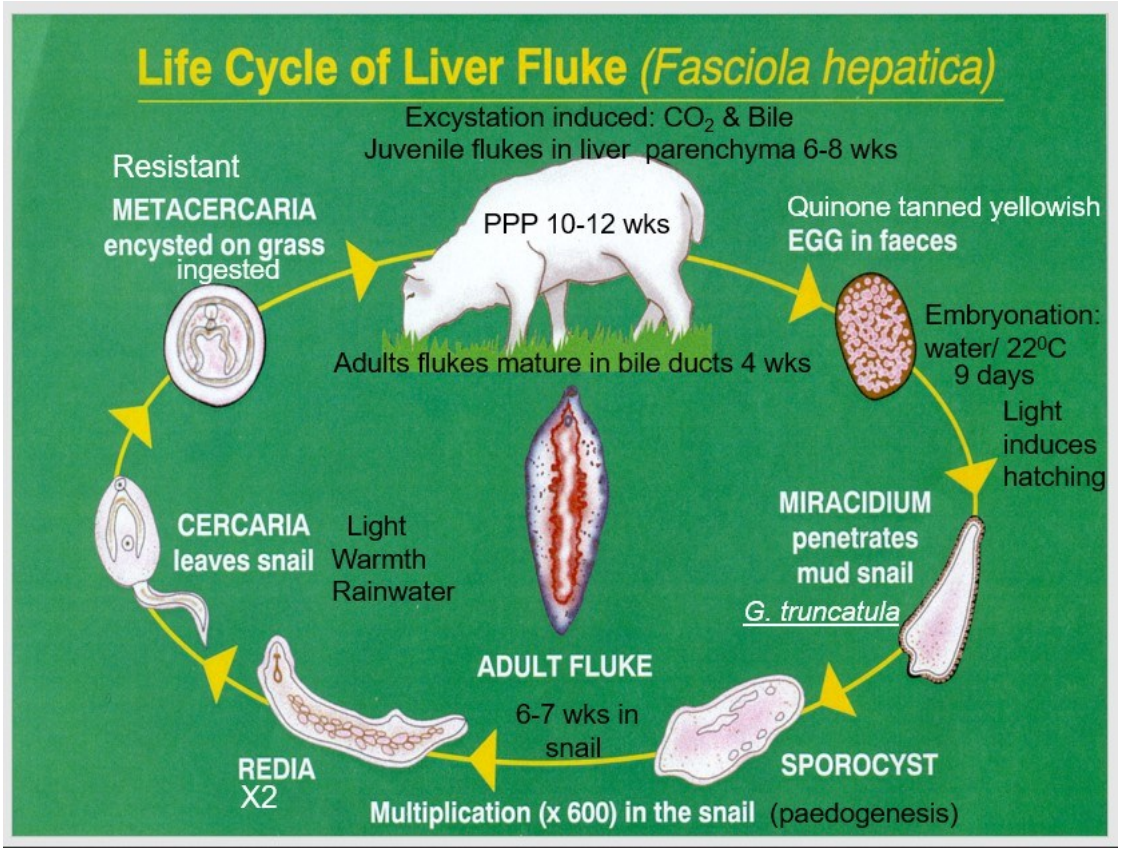


Figure 1.1: The life-cycle of *F.hepatica* (Taylor et al., 2016)

After about 6 weeks, rediae have developed into cercariae. Each redia produces 14 to 20 cercarial larvae, depending on the temperature (Abrous et al., 1999). Then, the cercariae loses its tail and develops a cyst around itself after releasing from the snails. The cercariae encysts on vegetation as metacercariae (Figure 1.2) which are about 0.2 mm in diameter, and it is in fact a juvenile fluke. If the metacercariae are formed in water, they can live for a year, but if they are formed on grass or vegetation then they survive only for a few weeks, they can withstand short periods of drying.

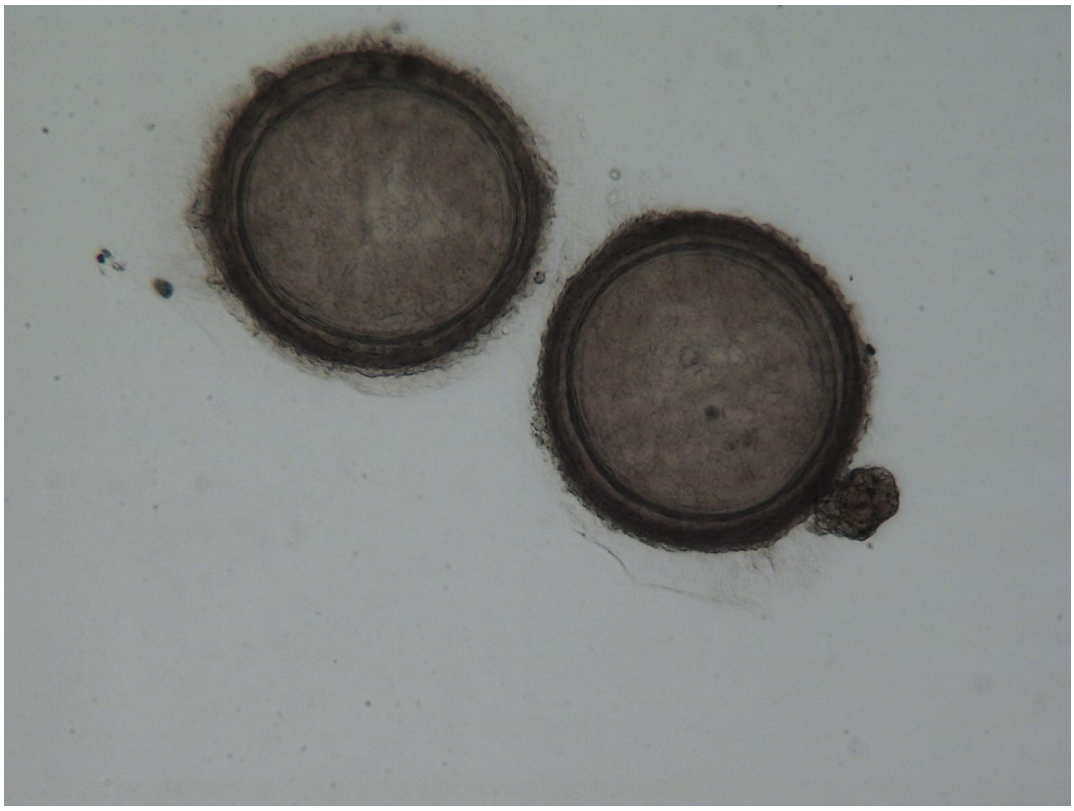


Figure 1.2: Metacercariae of *F. hepatica* (Photo by Mr.James McGoldrick)

The metacercarial cysts enter the body of animals through the metacercariae-contaminated grass while grazing. The metacercariae then hatch into juveniles, which burrow through the gut wall and migrate into the liver. Cattle and sheep can be infected by eating the contaminated grass or the consumption of metacercariae-contaminated water (Facey and Marsden, 1960; LaPook et al., 2000). Following a period of approximately six to eight weeks, the parasites migrate into the bile ducts where they develop into sexually mature adults (Figure 1.3), releasing 20,000 to 24,000 eggs per fluke per day (Boray, 1969). In humans, maturation from metacercariae into adult flukes takes approximately three to four months.

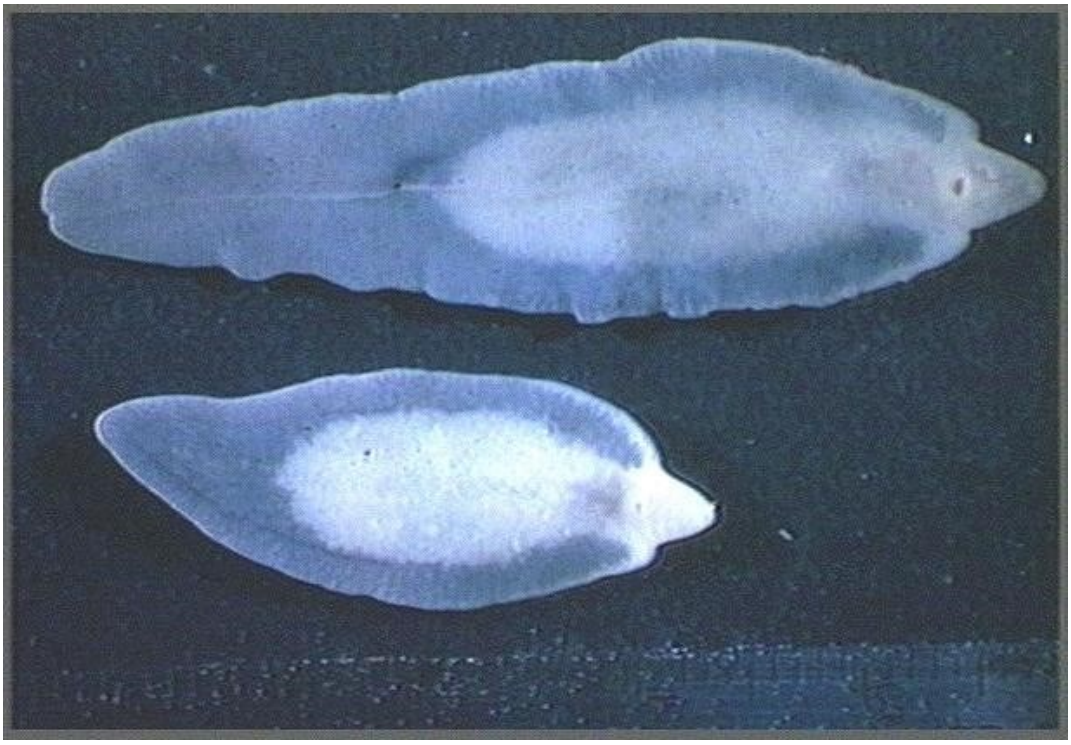


Figure 1.3: Adult *F. hepatica* (up to 30 mm by 13 mm) (Photo by Mr.James McGoldrick)

The disease caused by infection with *F. hepatica* is called fasciolosis, which causes liver damage and haemorrhage. Fasciolosis in sheep may be acute, subacute or chronic. The acute disease is the less common type of fasciolosis and occurs two to six weeks after the infection by over 2000 metacercariae. In the subacute disease, metacercariae are ingested over a longer period and while some have reached the bile ducts, where they cause a cholangitis, others are still migrating through the liver (Taylor et al., 2016). This form of the disease, occurring six to ten weeks after ingestion of approximately 500 to 1500 metacercariae, also appears in the late autumn and winter. Chronic fasciolosis, which is seen mainly in late winter/early spring, is the most common form of disease. It occurs four to five months after ingestion of about 200 to 500 of metacercariae (WHO, 1990; Abrous et al., 1999).

In Britain, the most suitable weather for *F. hepatica* reproduction is from May to October. A marked increase in numbers of metacercariae appear on pasture from August to October, before this the rediae develop in the snails during the summer and then the cercariae are shed from August until October. Alternatively, infections arise from the winter infection of snails and the metacercariae appear on the grass in May to June. This is called the winter infection. However, in most European countries, the summer infection of snails is more important and therefore the density of metacercariae in the grazing grass is increased during August to October.

1.2.2 Environmental effects on the life-cycle of *F. hepatica*

The life-cycle of *F. hepatica* is strongly affected by environmental factors. Climate and environmental factors are considered to be important elements to help us understand the occurrence of *F. hepatica* (Bennema et al., 2011). *G. truncatula*, the main immediate host of *F. hepatica*, is a species of air-breathing mud snail; and ponds or wetland are its main habitat. Climate change has been proved to have significant effects on the prevalence of fasciolosis (Kenyon et al., 2009).

1.2.2.1 Climate

Climate appears to be an important factor that can influence the whole life cycle of *F. hepatica*. It is long established that liver fluke infection risk is largely driven by the prevailing climatic conditions, particularly temperature and rainfall in UK (Met-office, b). The past ten to fifteen years in the UK have seen increasing fluke prevalence in livestock , as well as altered geographical distribution and seasonality of the disease (www.nadis.org.uk). These changes have been mainly attributed to changing climatic conditions, specifically warmer and wetter springs and summers and milder winters (Kenyon et al., 2009; Skuce et al., 2014). The related research shows that due to the climatic changes, an overall long-term risk level of fasciolosis will increase in all regions of the UK in the coming decades and also spatio-temporal variation in risk levels is expected (Fox et al., 2011).

Temperature is one of the most important factors affecting the life-cycle of *F. hepatica*. After its eggs are displaced in water with infected animals' feces, temperature and light are required (up to 22°C and 9-days light) for them to hatch into larvae. A mean day/night temperature of 10°C or higher is necessary both for snails to breed and for the development of *F. hepatica* within the snail, and all activity ceases at 5°C (Taylor et al., 2016). The development of cercarial larvae is also temperature- dependent. Environmental temperature is considered to be one of the most consistent factors influencing the development of *F. hepatica* (McCann et al., 2010a). Kantzoura's research shows that the geographical distribution of *F. hepatica* was affected by temperature and precipitation in south-eastern Europe (Kantzoura et al., 2011).

The whole life-cycle of *F. hepatica* occurs in wet surroundings. Water plays a key role in every stage of the life cycle. The ideal moisture conditions for snail breeding and the development of *F. hepatica* within snails are provided when rainfall exceeds transpiration, and the saturation of fields is attained (Taylor et al., 2016). The intermediate host *G. truncatula* is a snail, and the metacercariae in water can live for a year, but they can survive only for a few weeks on grass or vegetation (Taylor et al., 2016). The period of metacercarial occurrence affects the possibility of cattle infection, especially in the grazing areas. Geographical differences in surface water content can help to predict the spatial distribution of fasciolosis. Given a set temperature,

precipitation becomes a significant factor associated with the intensity of *F. hepatica* (Charlier et al., 2014b; Bosco et al., 2015). Indeed, heavy rainfall will promote the multiplication of *Fasciola* (Howell et al., 2015). Small water bodies, such as ponds, trenches and ditches, surrounded with grassland provide a suitable environment for *G. truncatula* and have been shown to provide suitable habitat for *G. truncatula* and have been shown to provide suitable habitat for *F. hepatica* (Vignolles et al., 2010; De Roeck et al., 2014; Charlier et al., 2014b). Also, water quality and ph are keys to *F. hepatica* (Charlier et al., 2014a; Manyangadze et al., 2016).

1.2.2.2 Soil moisture and soil material

McCann's research showed in some English and Walsh regions soil moisture has a positive impact on *F. hepatica* (McCann et al., 2010a). Soil texture influences land drainage and land surface, which can provide ideal habitats for *F. hepatica*. Poor quality land and soil with fine particles has poorer draining than soil with coarser particles. Soil with fine particles and flat poor quality land can form ponds, trenches or ditches. This will increase the risk level of *F. hepatica* on the soil with fine sand content. Also, soil ph has a negative impact on *F. hepatica* (Charlier et al., 2014b).

1.2.2.3 Landcover type and other topographical factors

Grazing strategies and herd management plays a major role in prevalence of economic infection levels of *F. hepatica* (Bennema et al., 2011). Improved grassland that is typically used for grazing will increase the risk level of fasciolosis if mowing strategies cannot be used (Charlier et al., 2014b). Grazing of boggy pastures and letting cattle access streams and ponds directly has been associated with higher risk (Howell et al., 2015). In addition, some research has suggested some plant species as a risk factor for *F. hepatica*; *Ranunculus* is a positive signal while reed-like plants can be considered as a less positive indicator (Charlier et al., 2016).

Topographical factors such as elevation, slope and aspect are also considered as factors which can affect the prevalence of *F. hepatica* because these factors will influence the temperature and moisture level within an area. A study undertaken in Iran shows that a positive correlation was obtained between altitude and *F. hepatica* (Ashrafi et al., 2015). Research found that the infection of fasciolosis decreased with elevation , which gave us an idea about how elevation effect *F. hepatica* (Lyngdoh et al., 2016).

1.3 Remote sensing and GIS approaches used in Parasitology

Earth Observation and Geographical Information Systems (GIS) approaches could be a good method to help with the identification of potential *F. hepatica* habitats. Moisture areas and small water bodies are ideal habitats for the intermediate host *G. truncatula*. In 1972, the first-generation earth resources technology satellite (later renamed as Landsat) was launched. Earth observation and remote sensing has been increasingly applied to a large variety of research fields. Earth Observation has proven to be a good way to detect and monitor the moist areas and dynamics of small water bodies such as ponds and ditches. Satellite imagery provides a stable and high-quality data source from which to analyse the prevalence of *G. truncatula*. GIS is a platform which can help researchers have a better understanding on the pattern of *Fasciolia* infection. Meanwhile, Meanwhile, GIS as a powerful tool can help farmers manage grazing strategies by mapping the spatial distribution of *F. hepatica*. A reasonable herb management strategy can reduce infection of *F. hepatica* (Bennema et al., 2011; Charlier et al., 2016). Combined Earth Observation with GIS tools can achieve economical and highly efficient agriculture management at the farm level.

Traditionally, vector habitats are mapped manually on ground surveys, which is time-consuming and also very expensive (Ozesmi and Bauer, 2002). Using Earth Observation and satellite images allows non-invasive, temporal monitoring in an automated way. Satellite images provide area-wide information making the traditional way of point-wise sampling of the Earth's surface more selective. Remote sensing images from space-borne sensors with resolutions from 1 km to 1m become more and more available at reasonable costs. For some remote sensing sensors already, large archives for periods of over 40 years are available via the World Wide Web (e.g. Landsat, NOAA-AVHRR, MODIS). This has resulted in the application of remote sensing in a large number of disciplines ranging from agriculture, environmental monitoring, and forestry to oceanography. It has been proved that remote sensing has significant advantages in monitoring inundation with *F. hepatica* (Alsdorf et al., 2007) and also can be an asset in parasitological research (Dambach et al., 2009; Vignolles et al., 2010). Recent research has shown the detection of vector habitats at a high spatial resolution in a large area in a semi-automated way and also indicated the importance of annual the *G. truncatula* life-cycle in the future studies of liver fluke (De Roeck et al., 2014).

Some studies combined the spatial distribution of *F. hepatica* with other factors when analysing potential infection areas, for example, management factors (Bennema et al., 2011) and characterised suitable small water bodies for *G. truncatula* (Charlier et al., 2014a). Some studies used the new methods to analyse the distribution pattern (Olsen et al., 2015) and used environmental factors, such as rainfall, temperature, soil moisture, to explain the spatial distribution pattern (McCann et al., 2010a). However, the application of remote sensing in this research in

F. hepatica remains undeveloped, especially the use of high-resolution satellite imagery. The most notable is DeRoeck's (2014) study which used object-based imaging analysis (OBIA) to present a classification of small water bodies based on 50 cm resolution Worldview2 imagery. Above all, cloud-free high-resolution images should be implemented in our research and the acquisition of images should occur at the time of seasonal peaks in vector abundance of both adults and juveniles. Many studies in this area have used field work associated with GIS analysis when assessing *F. hepatica* infection levels (Olsen et al., 2015) and manually mapped on aerial photography when classifying *G. truncatula* habitats (Halabisky, 2011). In recent years, due to the development of the satellite industry, the spatial resolution of microwave-sensed images has been improved significantly.

1.4 Aims and objectives

Earth Observation techniques and satellite imagery are widely used in the analysis of the land surface environment. The aim of this research is to use satellite data and GIS techniques associated with reference data (i.e. elevation, slope, soil moisture and soil temperature) to build a model for high-risk areas of *F. hepatica*. In addition, fieldwork is necessary as the reference data to correct the model.

The overall objective of this research is to develop an approach that can identify high risk areas for fasciolosis within a farm or paddock. The intended outcome of this research can be used for optimising management e.g. improve drainage, fence-off areas at high risk and minimising the use of anti-parasitic agents. To achieve the overall objective, farm-specific regression models between environment factors and *F. hepatica* were developed and applied.

Chapter 2

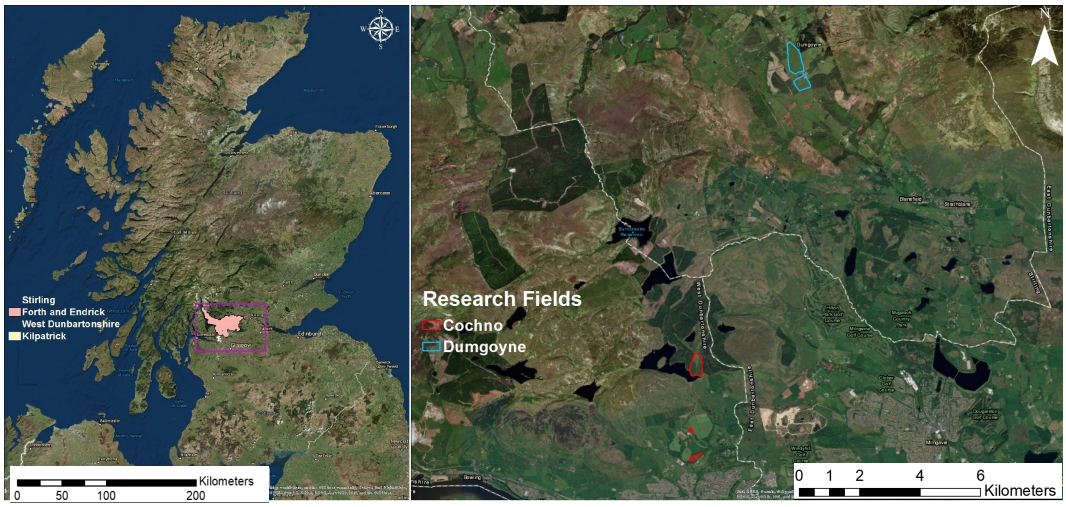
Methodology

2.1 Study Area

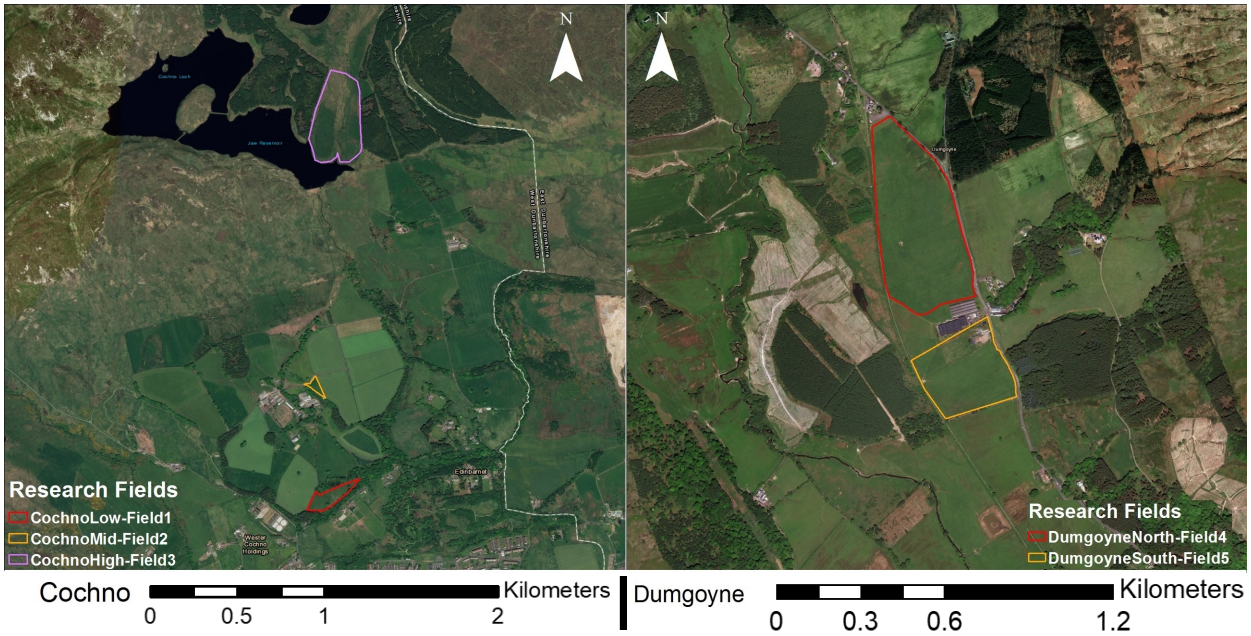
Two study areas just outside the city of Glasgow in Scotland; Cochno Farm and fields at Dumgoyne, were selected for analysis primarily because the study areas represented typical grazing fields, and also because of the availability of satellite imagery and aerial photographs, availability of weather data and soil type information, and their contrasting vegetation types. The Cochno study site is a University of Glasgow research farm and is used for core teaching and research activities at the School of Veterinary Medicine, and affiliated Research Institutes of the College of Medical, Veterinary and Life Sciences (www.gla.ac.uk/). The site covers an area of 344 ha, mainly consisting of agriculturally-improved grasslands and lands dominated by mosses or lichens. In this research, three fields were selected within the Cochno Farm (Figure 2.1). Selection of experimental fields was based on the variety of elevation and terrain, from lowland grassland to upland meadow. The Dumgoyne grassland is 13 miles away from Glasgow city and the area covers approximately 27 ha and is divided into two parts; the north field and the south field. In September 2017, the average daily precipitation of study fields was 5.83 mm per day and the average temperature was 12°C. The monthly average precipitation and temperature from January 2016 to September 2017 are given in Appendix A. The climate data were acquired from the Met Office website (www.metoffice.gov.uk/climate/uk) and it shows that in the summer of 2017 (from August to October), the study areas had more rainfall than the same period of 2016 and the monthly aver- age temperatures were about 12°C which was also slightly higher than 2016.

Table 2.1: Characteristics of the fields at the two study locations

	Fields	Highest Elevation (m)	Lowest Elevation (m)	Area (ha)
Cochno	Cochno Low	126.48	116.53	1.98
	Cochno Mid	173.02	160.94	0.46
	Cochno High	300.99	268.75	11.8
Dumgoyne	Dumgoyne North	33.29	20.75	18.2
	Dumgoyne South	31.28	23.21	8.45



(a) Two study areas



(b) Five research fields

Figure 2.1: Maps of study areas: (a) Two study areas; (b) Five research fields.

2.2 Earth Observation Imagery

This section describes the earth observation images and the methods of image processing used in this study. Images from three sensors were used as data sources. The main characteristics of each sensor and the reason for selection are presented below.

2.2.1 Sentinel-2

The Sentinel-2 images of 31 August 2017 was used in this study. Sentinel-2 is a wide-swath, high-resolution, multi-spectral imaging mission, supporting ESA Copernicus Land Monitoring studies, including the monitoring of vegetation, soil and water cover, as well as observation of inland waterways and coastal areas. The Sentinel-2 (https://earth.esa.int/web/sentinel/missions/sentinel-2) mission has the following capabilities:

- Multi-spectral data with 13 bands representing Top of Atmosphere (TOA) reflectances : four bands at 10 meters, six bands at 20 meters and three bands at 60 meters spatial resolution.
- Systematic global coverage of land surfaces from 56°S to 84°N
- Sun synchronous orbit at 786 km altitude, providing 14.3 revolutions per day
- The revisit frequency of each individual satellite is 10 days and the combined constellation revisit is 5 days.
- 290 km field of view
- free and open data policy

As shown in Table 2.2, four 10-m resolution bands, B2 (490 nm), B3 (560 nm), B4 (665 nm), B8 (842 nm) and one 20-m resolution bands, B11 (1610 nm) were used in this study. QA60 is a bit mask band which contains the cloud information: no cloud, dense cloud or cirrus cloud. All satellite images were collected and processed on the Google Earth Engine (GEE) online geospatial-processing platform. GEE is a cloud-based platform for planetary-scale environmental data analysis (https://code.earthengine.google.com). It provides an archive of publicly available remotely sensed imagery (Landsat, Sentinel-1/2) and other satellite data. GEE also provides tools and a code editor for rapid prototyping and visualisation of complex spatial analyses using a JavaScript API (Figure 2.2).

Table 2.2: Spectral bands for the SENTINEL-2 sensors(S2A) used in the study

Band number	Central Wavelength (nm)	Bandwidth (nm)	Spatial resolution (m)	Descriptions
B2	496.6	98	10	Blue
B3	560	45	10	Green
B4	664.5	38	10	Red
B8	835.1	145	10	NIR
B11	1,613.7	143	20	SWIR1
B12	2,202.4	242	20	SWIR2
QA60			60	Bitmask band

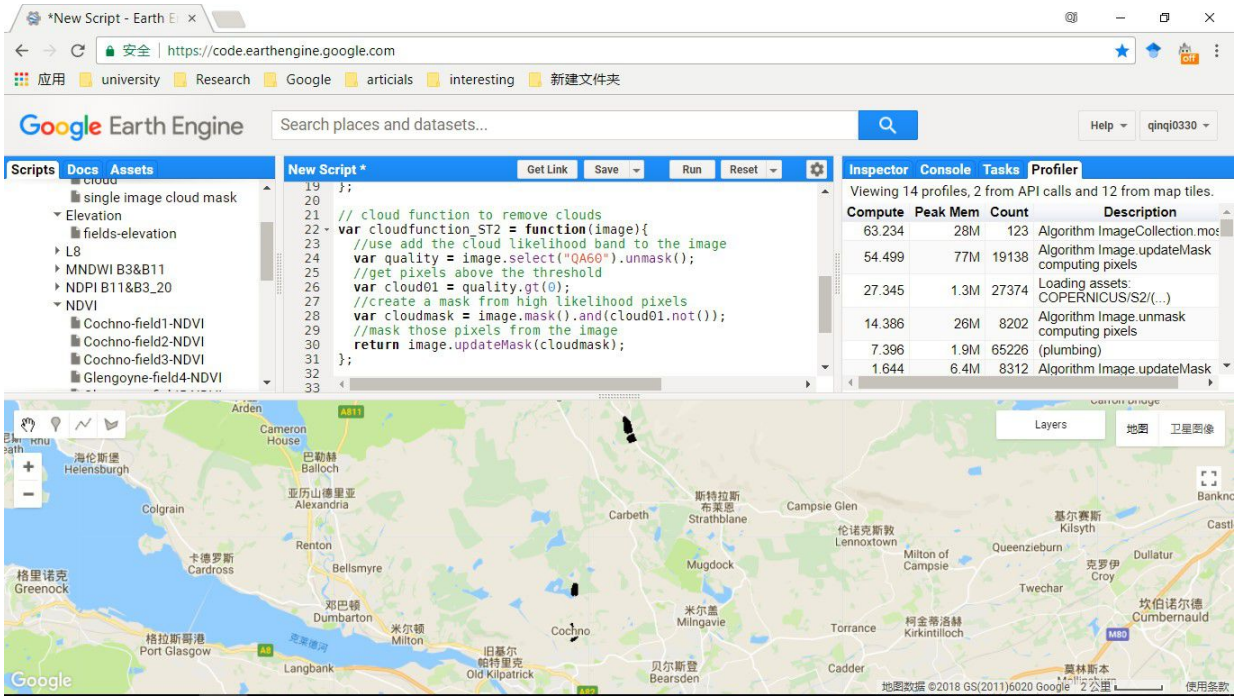


Figure 2.2: Google Earth Engine API.

2.2.1.1 Pre-processing of satellite images

The satellite imagery data acquisition process is affected by several factors that can decrease the quality of the images collected. This may have an impact on the accuracy of the information retrieved during the image analysis. In this research, all image analysis was completed on the GEE online platform using generated JavaScript codes. GEE also provides a large number of standard processing algorithms to help users access data easily and accurately. The pre-processing workflow is given in Figure 2.3. All JavaScript codes were given in Appendix E.

- Loading satellite images:

  Load the Sentinel-2 dataset using the Earth Engine collection ID (COPERNICUS/S2) and select image bounds and date period. The images produced by radiometric and geometric corrections and have been orthorectified. Display the images.

- Cloud-mask:

  To get clear cloud-free images of the research fields, we need to remove cloudy pixels. Sentinel-2 level-2C product provides the cloud mask information in Band QA60. The cloud mask is set as three values: 0 is a cloud-free pixel; 1 is a dense cloud pixel; 2 is a cirrus cloud pixel (Sentinel-2 user handbook, 2015). In this study, we just need the cloud-free pixels with a QA Band value equal to zero.

- Resampling:

  The Sentinel-2 product has three spatial resolutions (10 m, 20 m and 60 m). In this research, the band B2, B3, B4 and B8 are 10-m resolution, which were down-scaled to 20 m, same as B11 and B12. The Google Earth Engine provides the algorithms to resample images with nearest neighbour.

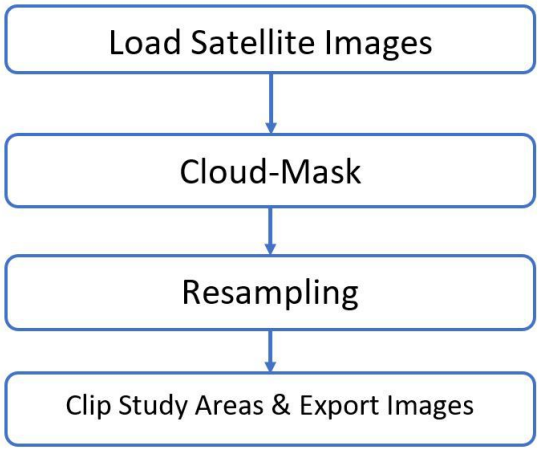


Figure 2.3: Workflow of pre-processing of satellite images.

2.2.2 LiDAR

Light Detection and Ranging (LiDAR) is an airborne mapping technique, which uses a laser to measure the distance between the aircraft and the ground. Up to 300,000 measurements per second are made of the ground, allowing highly detailed terrain models to be generated at spatial resolutions of between 25 cm and 2 m. LiDAR data can represent a very accurate terrain model of the earth surface with ground cover features and detailed geographic information. LiDAR is an option in remote sensing technology that optimises the precision of biophysical measurements and extends spatial analysis into the third dimension (Popescu, 2007). LiDAR data of Scotland are published by the UK Environment Agency through the Digimap online platform (http://digimap.edina.ac.uk/lidar). The Environment Agency's LiDAR data archive contains digital elevation data derived from surveys carried out by the Environment Agency's specialist remote sensing team. Data are available in Digital Surface Models (DSM) and Digital Terrain

Models (DTM) with at 2 m, 1 m, 50 cm, and 25 cm resolutions across the UK.

The 2015 Scotland LiDAR data were used in this research and were received as Digital Terrain Models (DTM) at 1 m resolution. To cover all five research fields, twelve 5×5 km images tiles were collected (Figure 2.4) and the images were received in ascii data format.

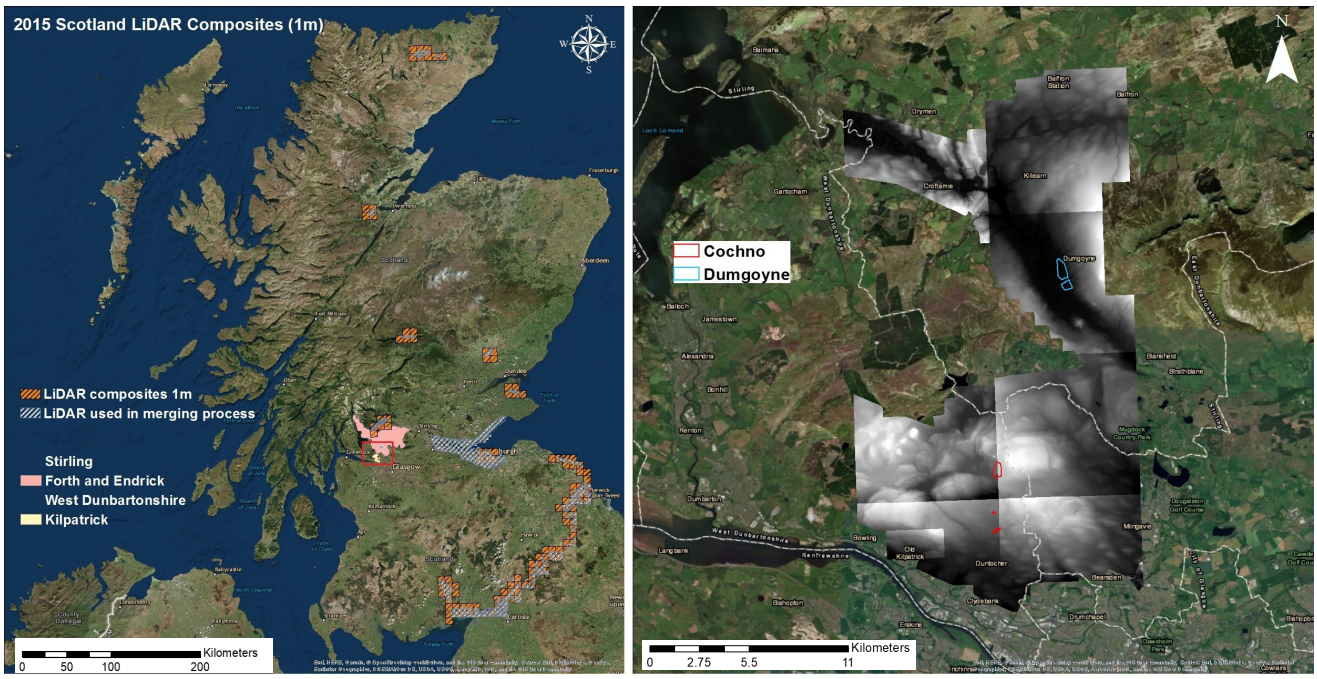


Figure 2.4: LiDAR images: (a) Scotland 1 m LiDAR composites; (b) LiDAR tiles used in this study.

2.2.3 Aerial photography

Aerial photography has significant advantages in detection of detailed topographical information because of its minimum mapping unit size. Additionally, low-level photography, using helicopters or unmanned aerial vehicles (UAVS) can provide even smaller minimum measurement units (Hackney and Clayton, 2015) and therefore is commonly used in the application of accurate field surveys and small field survey. Aerial photographs at 25 cm resolution were obtained for this research from the Digimap online platform (http://digimap.edina.ac.uk/aerial). Digimap provides free access to aerial photographs in JPG data format. This research used 68 aerial photographs acquired in 23 August 2015 (Figure 2.5).

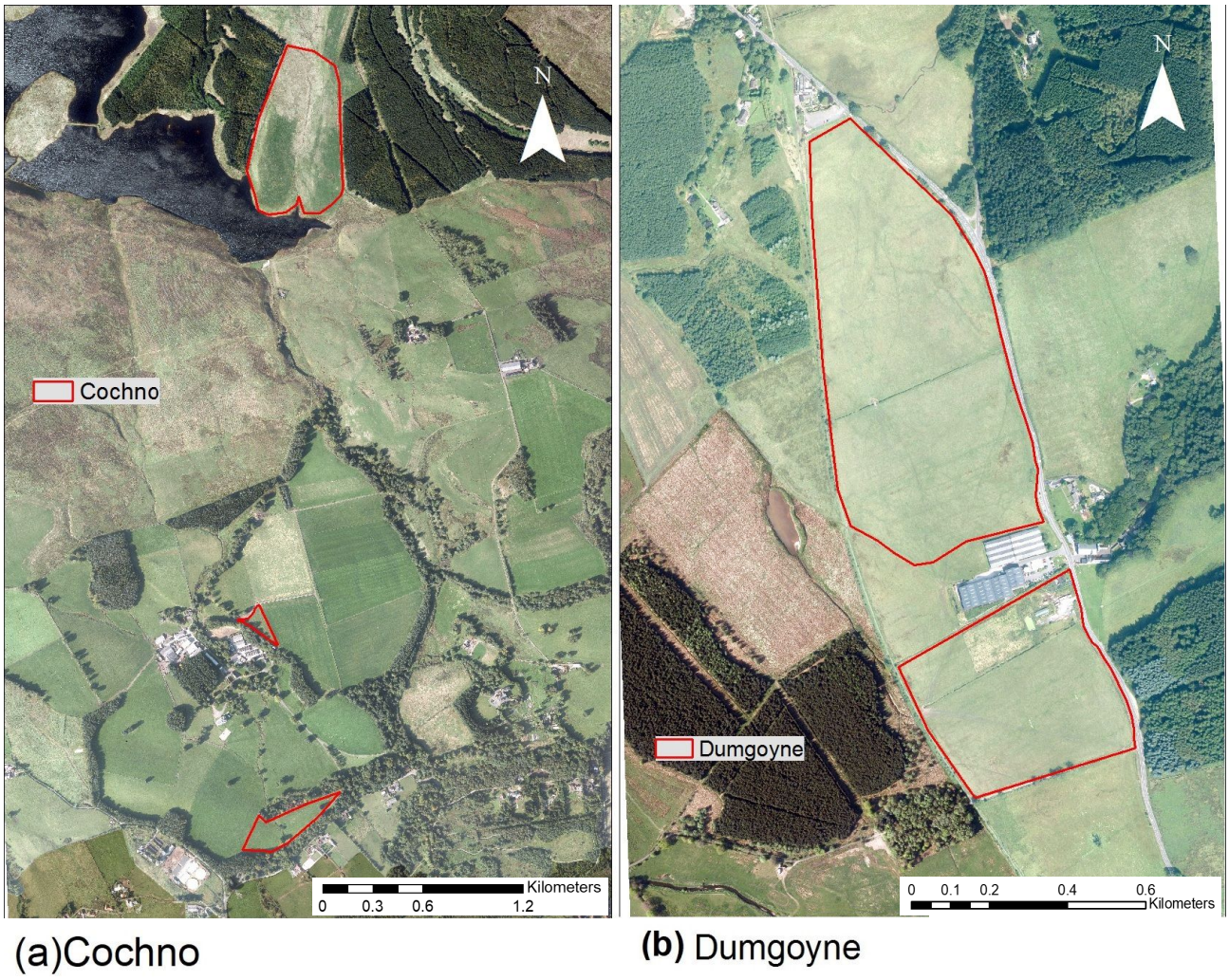


Figure 2.5: Aerial photography: (a) Cochno; (b) Dumgoyne.

2.3 Reference data

2.3.1 Terrain data

Environmental factors have significant influence on the life cycle of *F. hepatica* and, therefore are considered as reliable predictors of the risk of fasciolosis. The environmental and geographical data are major sources to help understand the geographical distribution of *F. hepatica* and the prevalence of fasciolosis (McCann et al., 2010b; Manyangadze et al., 2016). In this research, ArcMap Desktop 10.3 was used to store, manage, analyse and visualise the datasets. ArcMap Desktop is a powerful and professional GIS application used for all map- based tasks, including cartography, map analysis, and editing. It offers different tools that can perform any GIS task,from simple to advanced, including mapping, geographic analysis, data editing and compilation, data management, visualisation, and geoprocessing.

Topographical information such as elevation, slope and aspect have impact on *F. hepatica* because they affect weather and vegetation (McCann et al., 2010b; Bennema et al., 2011). Elevation data is represented in the LiDAR DTM images while slope and aspect can be calculated from the LiDAR DTM. The LiDAR DTM images were based on the British National Grid and needed to be re-projected to the UTM/WGS84 in an initial step. The steps of processing LiDAR DTM data are shown in Figure 2.6.

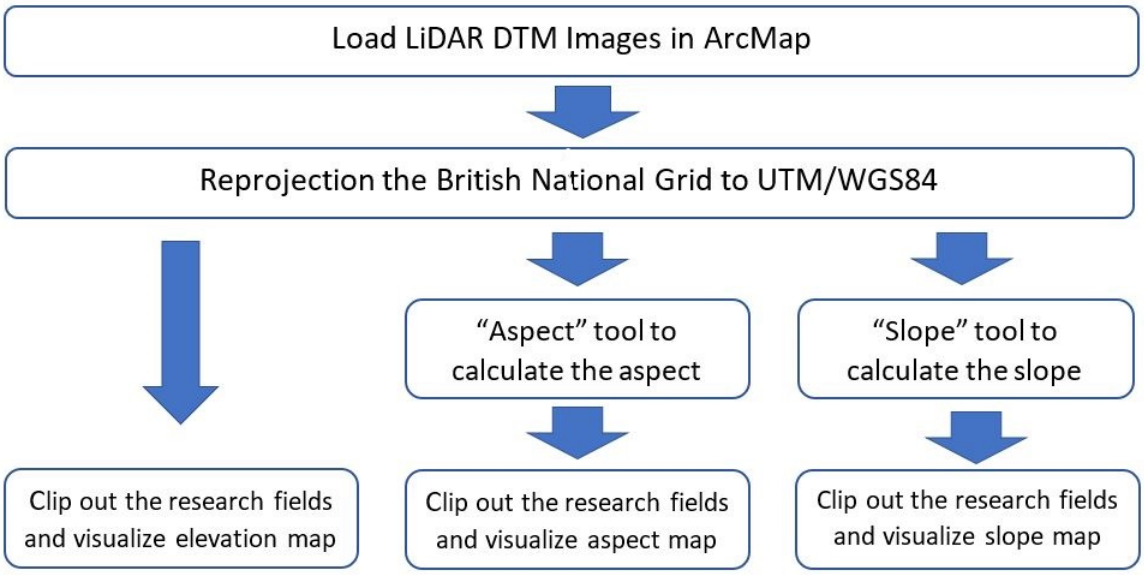


Figure 2.6: Workflow of pre-processing of LiDAR DTM Data.

ArcMap Desktop was used to extract the elevation data from the LiDAR DTM images. The elevation was based on the WGS84. The elevation map of research fields is shown in Figure **??**.

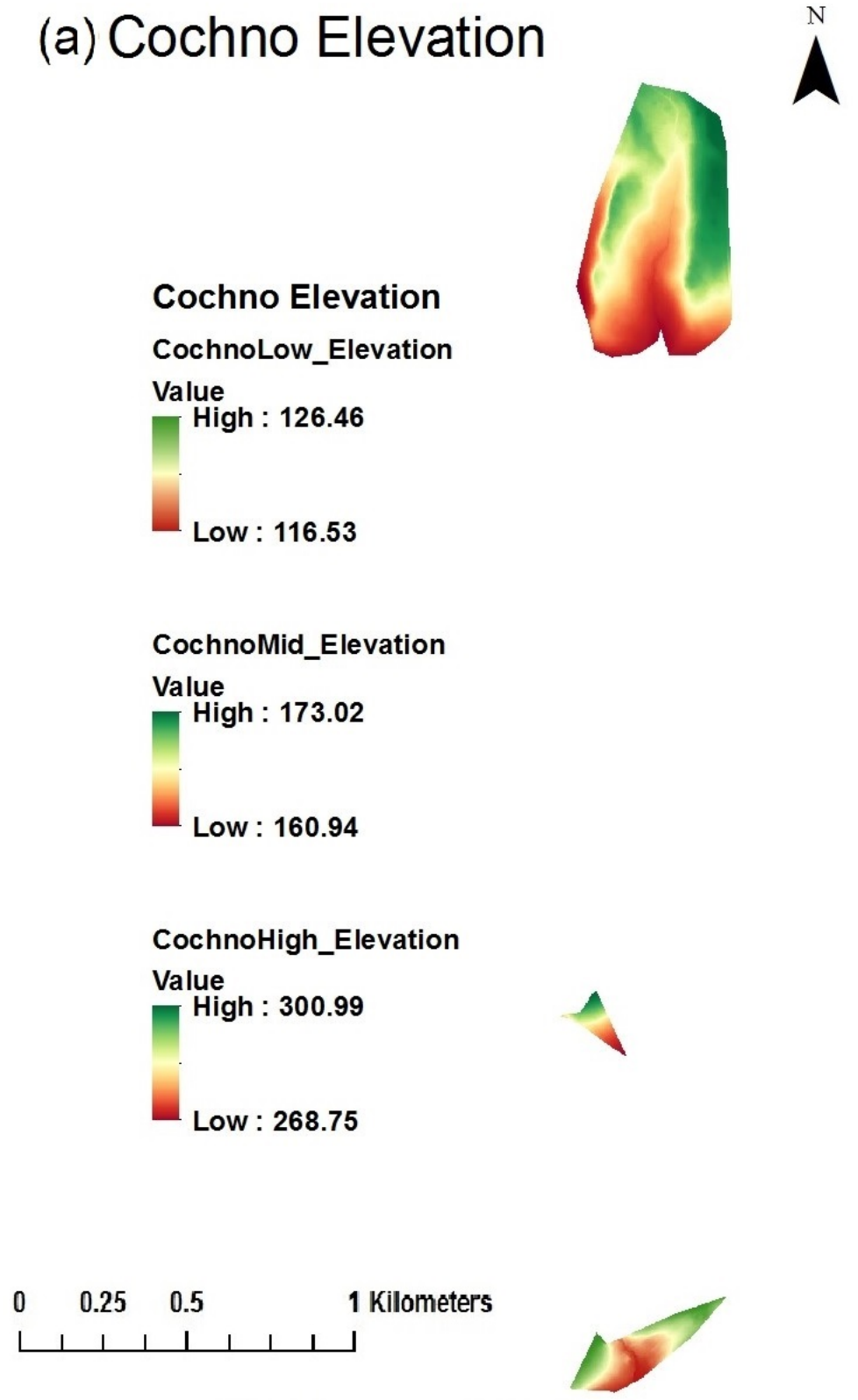
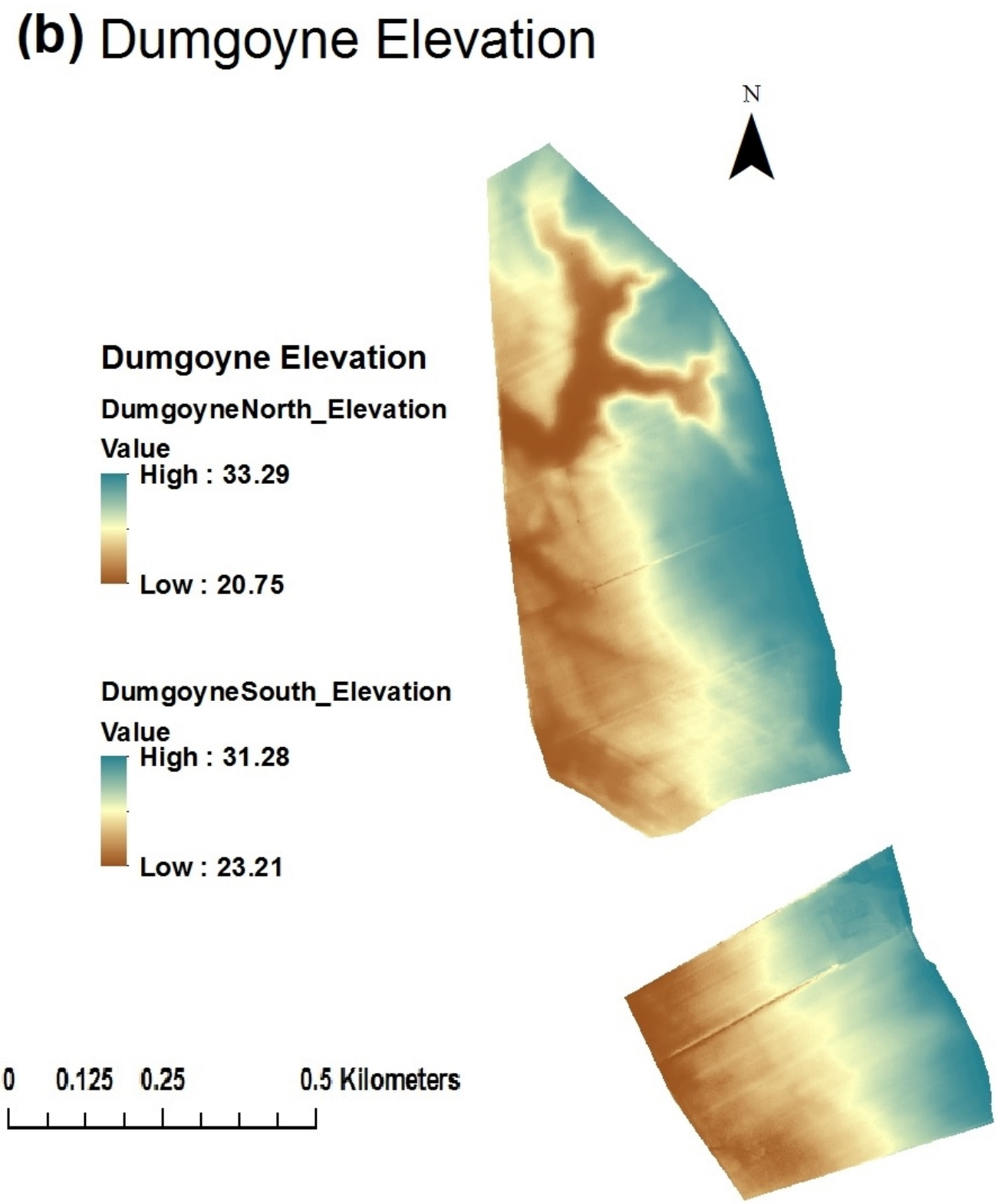


Figure 2.7: Elevation maps of research areas: (a) Cochno; (b) Dumgoyne.



The Slope tool in ArcMap was used to calculated slope of the research fields. The Slope tool identifies the steepness at each cell. Slopes were measured in units of degrees, using the following 2.1, where the $[\delta z/\delta x]$ means the rates of change of the surface in the horizontal and the $[\delta z/\delta y]$ means the rates of change in the vertical directions.The number 57.29578 is a truncated

version of the result from 180/pi.

$$slope\_degrees = \arctan\left(\sqrt{[\delta z/\delta x]^2 + [\delta z/\delta y]^2)}\right) * 57.29578 \tag{2.1}$$

The lower the slope value, the flatter the terrain; the higher the slope value, the steeper the terrain. The slope map of research fields is shown in Figure 2.8.

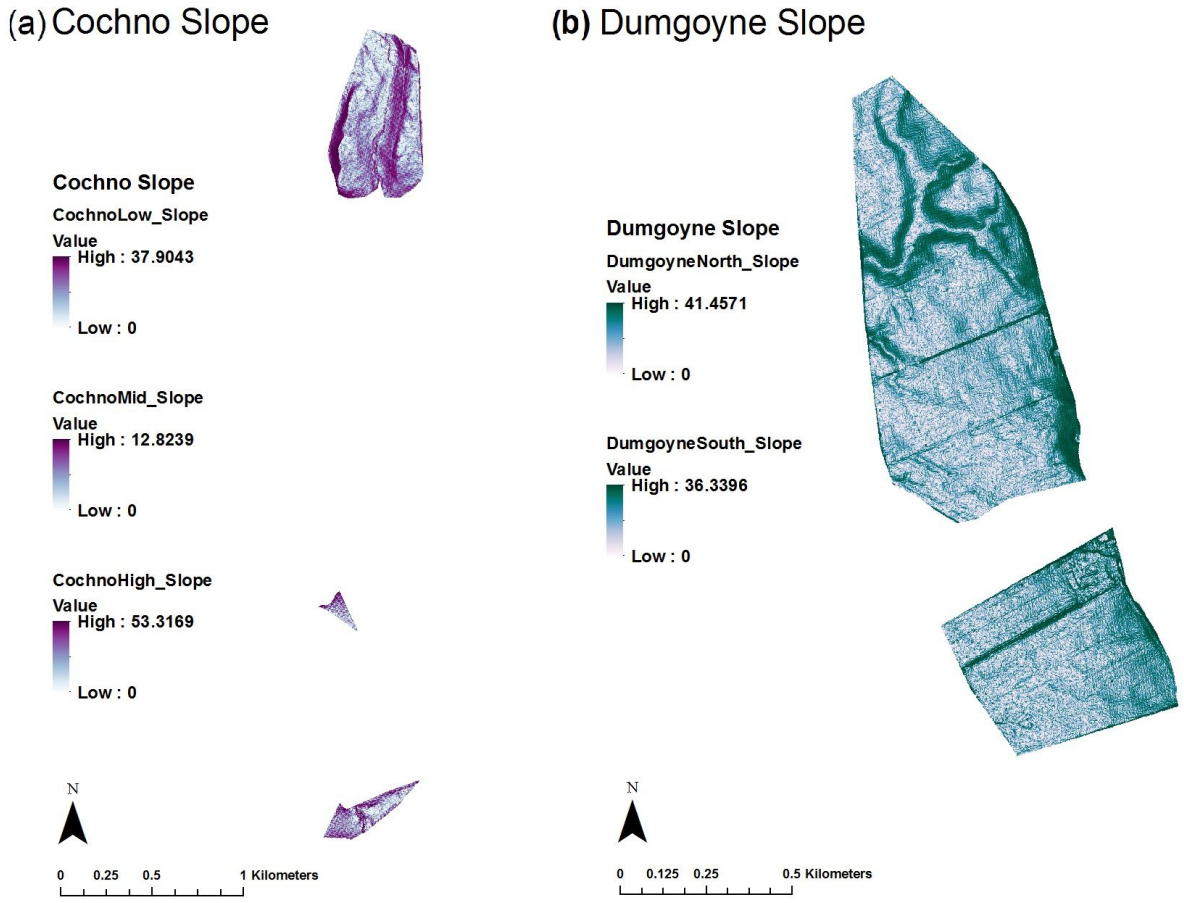


Figure 2.8: Slope maps of research areas: (a) Cochno;(b) Dumgoyne.

To calculate the aspect, the Aspect tool in ArcMap was used. The Aspect tool identifies the direction in which the downhill slope faces. The value of each cell in the output raster indicates the compass direction the surface faces at that location. It is measured clockwise in degrees from 0 (due north) to 360 (again due north), coming full circle. Flat areas having no down-slope direction are given a value of -1. The Figure 2.9 below shows the aspect of five research fields.

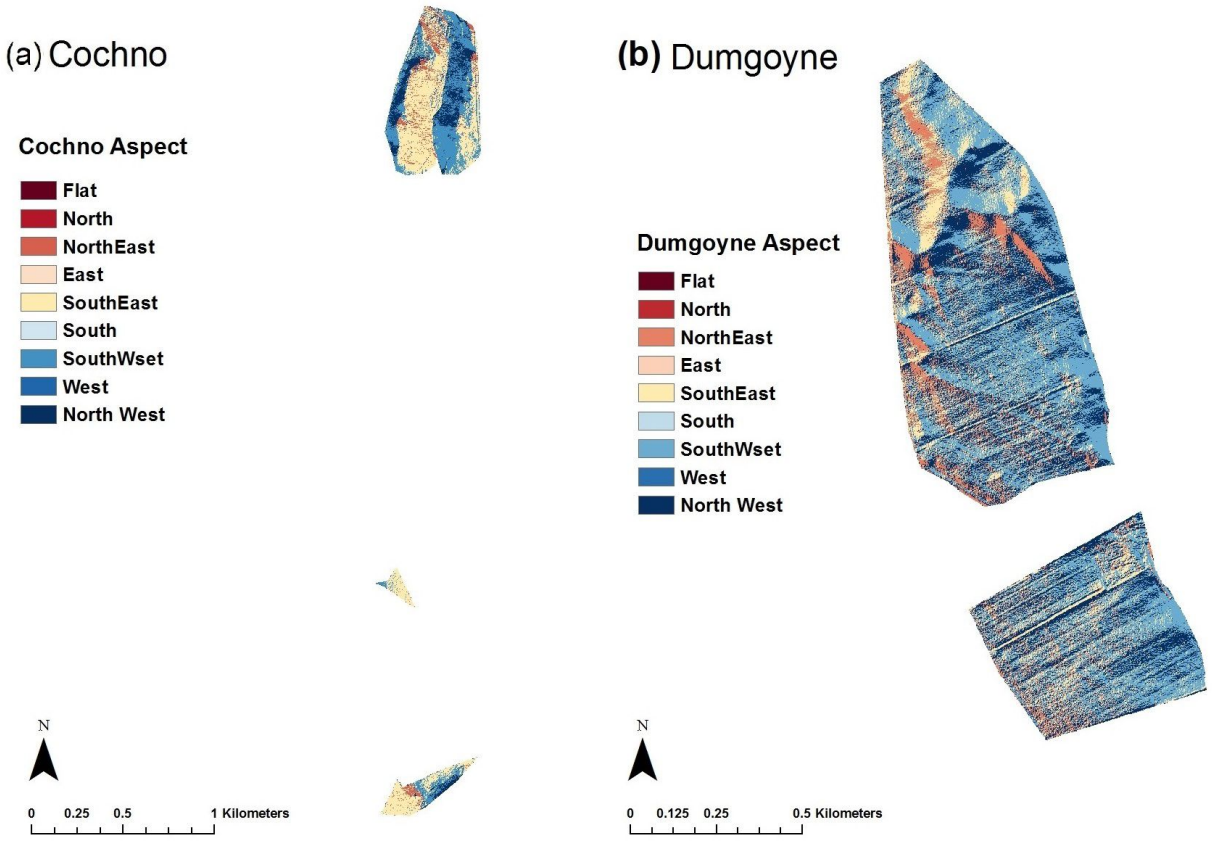


Figure 2.9: Aspect maps of research areas: (a) Cochno; (b) Dumgoyne.

2.3.2 Soil texture

Soil texture and soil particles may influence the presence of liver fluke and its host. Poor quality land with fine sand content is more suitable for the fluke's intermediate hosts, *G. truncatula*, because soil with fine particles has poorer drainage than soil with coarser particles (McCann et al., 2010b).The Soil-Parent Material (PM) Database was used in this research. The Soil-Parent Material (PM) Database was published by the British Geological Survey and can be downloaded from the Digimap online platform. The Soil-Parent Material (PM) Database is part of a series of GIS maps designed to help environmental scientists and consultants assess the characteristics of the 'near-surface' weathered zone. A parent material is a soil-science name for a weathered rock or deposit from, and within which a soil has formed. In the UK, parent materials provide the foundations and building blocks of the soil, influencing their texture, structure, drainage and chemistry (Avery et al., 1980).

The data were obtained in Shapefile format in the British National Grid coordinate system. Soil-Parent Material data using in the study were acquired from the Digimap (https://digimap.edina.ac.uk). The soil data are in the 1:50,000 scale and contain over 30 rock and sediment characteristics built upon the standard DiGMapGB-50 geological dataset, which offers 44 attributes to describe soil

properties. As soil texture and soil particles are key parameters of the life cycle of *F. hepatica*, this research selected four attributes which represent soil texture and soil particles.

- SOIL_TEXT and EST_TEXT
  Soil_text and Est_text text provides a general classification of soil texture from measured samples of soils overlying this parent material. Soil texture classes are based on a UK classification of soil texture designed by The National Soil Research Institute (Hodgson, 1974).

- Grain-size Name and Particle Diameter
  A geological description and qualitative classification of the grain sizes expected for this parent material based upon the information in the minimum, maximum and dominant grain size fields.

2.4 Remote sensing indices

In remote sensing, radiometric or spectral ratios are enhancement techniques where raster pixels from one spectral band are divided by corresponding values in another band. Both above indicators have the same form of function; the choice of frequency bands used makes them suitable for specific purposes.

2.4.1 Normalized difference water index

The Normalized Difference Water Index (NDWI) is one of the most widely-used methods to indicate water or moisture conditions. Surface water and the water content of land features can be reflected on NDWI values. NDWI is shown by a formulation which uses the combination of different visible light wavelengths. In 1996, Gao proposed a formulation using near-infrared (860 nm) and short-wave infrared (1240 nm) wavelengths to monitor changes in water content of leaves (Gao, 1996). The Short Wavelength Infrared (SWIR) reflects changes in vegetation canopy water content and sponge structure. The near-infrared is affected by the internal structure of the leaves and the dry matter content of the leaves but is not affected by the moisture content. The combination of NIR and SWIR eliminates the changes caused by leaf internal structure and leaf dry matter content and improves the accuracy of restoration of vegetation moisture content (Ceccato et al., 2001). The NDWI is calculated as follows:

$$NDWI = \frac{(nir - swir)}{(nir + swir)} \tag{2.2}$$

In this study, *nir* stands for the spectral reflectances measurements acquired in B8 (824 nm) and *swir* stand for the spectral reflectance measurements acquired in B11 (1610 nm). These spectral reflectances are themselves ratios of the reflected incoming radiation in each spectral band individually, hence they take on values between 0.0 and 1.0. By design, the NDWI itself thus varies between -1.0 and +1.0. NDWI values between -1.0 and 0 represents bright surface with no vegetation or water content, while values above 0 to +1.0 represents water content.

2.4.2 Normalized difference vegetation index

Calculation of vegetation indices is one of the most common transformations applied to image data. It serves to highlight subtle variations in the spectral responses of various surface covers. A vegetation index can be used to estimate continuous variables like, e.g. LAI and biomass. Normalized difference vegetation index (NDVI) is one of these vegetation indexes. NDVI was one of the most successful of many attempts to simply and quickly identify vegetated areas and their condition, and it remains the most well-known and used index to detect live green plant canopies in multispectral remote sensing data (Pettorelli, 2013). The NDVI is calculated from these individual measurements as follows:

$$NDVI = \frac{(nir - red)}{(nir + red)} \tag{2.3}$$

In this study, *nir* stands for the spectral reflectances measurements acquired in B8 (824 nm) and *red* stands for the spectral reflectances measurements acquired in B4 (665 nm). Visual or digital interpretation of the output image/raster created is similar to NDWI; the NDVI itself thus varies between -1.0 and +1.0. NDVI values between -1.0 and 0 represents earth surface with poor vegetation or bare soil, while values above 0 to +1.0 represents sufficient vegetation.

2.5 Preliminary risk map

Before field survey and sampling, a preliminary risk map of *F. hepatica* was generated based on the NDWI, NDVI, slope and soil texture. The risk map showed a direction to the potential high level of *F. hepatica* contamination of pastures and therefore was used as a reference when selecting the sampling locations. In the following field survey and laboratory analysis, the accuracy of the risk map was examined. The NDWI and NDVI layer were sourced from Sentinel-2 satellite images and the slope and soil were obtained from the Digimap online platform, described in the previous sections.

The NDVI, NDWI and Slope values were classified by five classes. For the NDVI and NDWI,

the higher values corresponds to higher risks (Bennema et al., 2009) and for the slope, the flat terrain may have the higher risks (Tum et al., 2004). As for the soil layer, the risk class was identified by the soil texture, basically soil with fine particles more poor drainage than soil with coarser particles and therefore a higher risk level. This study assumes that each classification was equally distributed. The classification was carried out on ArcMap 10.3 using the tool *Reclassify*. The reclassification method was equal interval. The risk layers of NDVI, NDWI, slope and soil texture were added up to create a risk map for each study area. All risk maps are presented in Chapter 3.

2.6 Field survey and laboratory analysis

2.6.1 Field survey method

2.6.1.1 Selection of sampling points

In this research, we selected 12 sampling points for each of five fields, based on the preliminary risk map and the actual field situation. Sampling points were selected randomly using ArcMap Desktop tool 'Create Random Points'. All 60 samples were classified by research fields and numbered according to the rule ' ***research fields + number***'. The latitude and longitude of each of the points were recorded in a hand-held GPS device under the WGS84. The coordinates of sampling points are listed in Appendix B.

2.6.1.2 Materials

Main materials used in the field survey are listed below:

- GPS navigation device
  Garmin GPX60 with the navigation accuracy between 5 to 10 meters (https://support.garmin.com/e
  GB), recording latitudes and longitudes in WGS84 cordinate system.

- Soil Sensor: Stevens HydraProbe
  A soil sensor HydraProbe was used in this study to measure soil moisture and soil temperature. HydraProbe provides the measurement of Soil Moisture in units of water fraction by volume (WFV), with the ± 0.01 WFV accuracy for most soils and ± 0.03 max for fine textured soils. The soil temperature measurement is within $\pm 0.3^{\circ}$C accuracy and the unit of soil temperature is $^{\circ}$C. (`https://www.stevenswater.com/products/hydraprobe/`). The accuracy may vary with some soil textures.

- Other materials
  A 50×50 cm quadrat, wood stakes and shears

2.6.1.3 Sampling approach

The metacercarial cysts are formed on grass or vegetation and , therefore, the main activity of the field survey was to cut and remove grass at specified locations and quantify the metacercariae from standardised areas of grass and from standardised weights of grass. We also measured soil moisture and soil temperature at each sampling location. The survey at each sampling point used quadrat sampling methods. In the Dumgoyne sites, the survey was conducted on 5 September 2017: twenty-four quadrat samples were taken in total. The survey in the second study area at Cochno was conducted on 6 September 2017, with a total of thirty-six twenty-four quadrats recorded.

First, following the GPS device, the location of each pre-selected sampling location was identified and the quadrat was placed. Within each quadrat, grass height was randomly measured five times using a ruler; and three points were chosen to measure soil moisture and soil temperature. The measurements of grass height were recorded by hand and the measurements of soil were recorded in the soil sensor. The grass was cut down to the ground within each quadrat. Measurements of soil moisture (%) and soil temperature (°C) were made in each quadrat using the soil sensor. Latitude and longitude co-ordinates were input to the ArcMap desktop to show the position of each sampling point on the study area map. The average value of the three measurements at each sampling point was taken as the measurement result at that point.

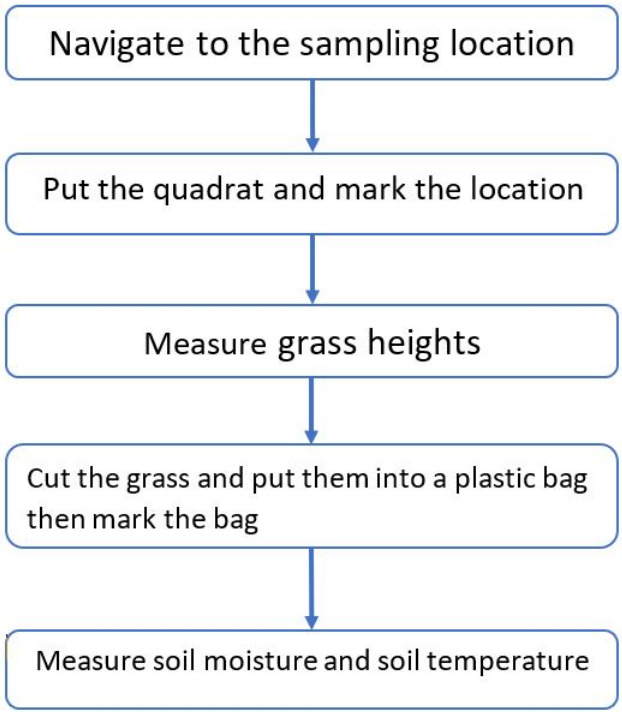
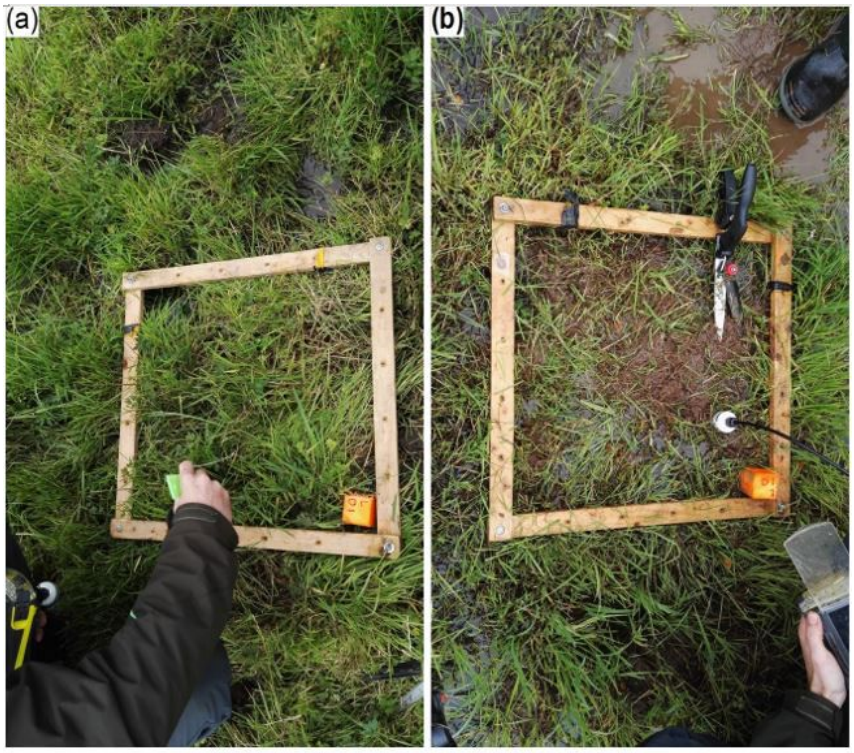


Figure 2.10: Workflow of field survey and quadrat sampling.

(a) Put a quadrat on the ground (e.g. N05); (b) Mark the location with a wooden stick and colour it in orange (e.g. N12)



(a) Measure the grass heights (e.g. L01); (b) Measure the soil moisture and soil temperature (e.g. L0)

(a) Grazing the grass within the quadrat (e.g. N01 after grazing); (b) A tub for sheep drinking water in the Dumgoyne South

Figure 2.11: Field survey

2.6.2 Laboratory analysis

The metacercarial cysts of *F. hepatica* within pasture grass of sampling locations were counted in the laboratory. The procedure of laboratory analysis followed the Manual of Veterinary Parasitological Laboratory Techniques reference book, established by the UK Ministry of Agriculture, Fisheries and Food (MAFF) in 1986 (Great Britain. Ministry of Agriculture and Food, 1986).

2.6.2.1 Materials

All materials used in the lab work were listed as followed, including the lab equipment and the chemical solution used to store samples.

- Laboratory balance
- Washing Machine
- 10% Formalin solution
- 38 μm Laboratory Test Sieves (ENDECOTTS)
- Stereo-Microscopes: eyepiece $\times 10$ and objective $\times 10$, $\times 20$ and $\times 50$
- 1 ml air-displacement pipette

2.6.2.2 Working Procedures

There are 60 samples in this research. For each sample, the following working procedure was followed. The work flow was shown in Figure 2.12.

(i) First, grass weight was measured weight of each grass sample was calculated as the weight of the sample minus the weight of empty plastic bag (22 g in this study).

(ii) The metacercarial cysts are normally found on and vegetation and are very hydrophilic. Washing grass samples moderately can make metacercarial cysts into water. Grass was first washed in a washing machine with about two litres warm water for nine minutes. Metacercarial cysts are about 0.2 mm in diameter and grass was first wrung out into the washing and then remove grass out of the washing machine and placed grass back in the transport bag. Keep the water and pull into a clean bucket.

(iii) The liquid in the washing machine was then removed from the washing machine and placed into a bucket. The washing machine was then rinsed with water 2 or 3 times and the wash retained in the bucket. Then the liquid was poured through a 38 μm sieve, and the filtrate in the sieve was retained. The sieve was then washed and the process repeated a further two or three times. until the filtrate was less than 75 ml. The filtrate was placed into a new plastic container and 25 ml formalin was added. Rinse the washing machine and buckets with clear water for the next sample.

(iv) Samples were then shaken 10 times and 1 ml was pipetted into a petri dish, to which a small amount of water was added to cover the entire dish before examination under the microscope using $\times 10$ objective for screening with $\times 20$ or $\times 50$ objectives used for confirmation. Seven replicate samples were examined from each sample of grass. The metacercarial cysts have some noticeable characteristic features: firstly, shape is a perfect circle of about 0.2 mm diameter, which is quite big. Secondly, the metacercarial cyst of *F. hepatica* has an obvious wall, which creates a noticeable dark edge under the microscope and it is very clear within the metacercariae sometimes with small bubbles.

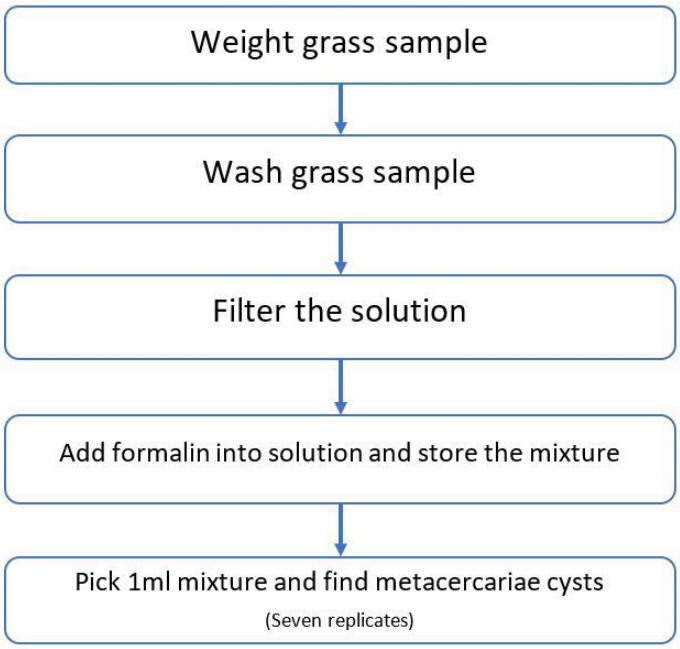


Figure 2.12: Workflow of laboratory analysis.

2.6.3 Calculation of metacercarial cyst count

We took 7 replicates of the metacercarial cyst counts for each sample, which means the number of metacercarial cysts in 100 ml solution could be calculated as the following equation :

$$Metacercariae = (n/7) * 100 \tag{2.4}$$

In which n is the number of metacercarial cysts counted in each replicate. The number of metacercarial cysts per gram of forage (Yield) was calculated by using the equation:

$$Yield = Metacercaria/GrassWeight \tag{2.5}$$

Yield reflects the contamination of metacercarial cysts in every gram of forage grass. To use in the logistic models, the Yield values were transformed to Yield2 using the equation:

$$Yield2 = Yield * 0.1 \tag{2.6}$$

The Yield2 values means the number of metacercariae per 10 g. It follows the same distribution as the Yield.

2.7 Statistical analysis methods

This section describes the statistical methods used in assessing the relationships among the variables as well as the variation among fields. All methods were carried out using the software R 3.5.0 (R Core Team 2016). The confidence interval used in this study is 95%. According to the descriptions in the previous sections of this chapter, there were 58 samples and for each sample, 11 variables were used in this study. All variables were numeric and recorded to the third decimal. The entire data are given in Appendix C. The 11 variables were divided into five datasets:

- Dataset 1: *F. hepatica* density in pasture
  Dataset 1 contained two variables: (1) Metacercarial cysts: the number of metacercarial cysts; (2) Yield: the number of metacercariae per gram

- Dataset 2: Remote sensing indices
  Dataset 2 contained two variables: (1) NDWI: normalised difference water index; (2) NDVI: normalised difference vegetation index. All values obtained from the NDWI and NDVI maps of each fields on 31 August 2017 based on the coordinates of each sampling points.

- Dataset 3: Topographical factors
  Dataset 3 contained three variables: (1) Elevation: the altitude of each sampling point; (2) Slope: the slope of each sampling point; (3) Aspect: the aspect of each sampling point. All values were sourced from the DEM images.

- Dataset 4: Soil properties
  Dataset 4 contained two variables: (1) Soil Moisture: soil moisture of each sampling point; (2) Soil Temperature: soil temperature of each sampling point. Soil moisture and temperature were measured from the field survey using a Stevens HydraProbe soil sensor.

- Dataset 5: Pasture
  Dataset 5 contained two variables: (1) Grass Height: the height of sampling grass; (2) Grass Weight: the wet weight of sampling grass, which were measured results in field survey and laboratory analysis.

2.7.1 Preliminary analysis

The results of the preliminary analysis are presented in Chapter 4 and make use of summary statistics and graphical tools such as histograms and kernel density curve (Gaussian), which were produced in the statistical software package R 3.5.0 (R Core Team, 2018). The process was particularly crucial for the exploratory statistical assessments. Before further statistical tests,

some preliminary tests were required to make sure that the relevant test assumptions are met. Shapiro-Wilk normality test (Shapiro and Wilk, 1965) were used in this study, and the R 3.5.0 function *shapiro.test()* was applied to assess the normality of data. If the p-value given by the Shapiro-Wilk test was greater than 0.05, we conclude that the data follow a normal distribution. If the data do not follow a normal distribution, the data are required to be transformed into a normal distribution to reduce the skewness in the data. The Box-Cox transformation was carried out for highly skewed data. The Box-Cox procedure is available with the *boxcox()* function in the MASS package in R 3.5.0. After transformation, a Shapiro-Wilk test was applied once again. If the new data are still not a normal distribution, non-parametric tests will be used. The results of preliminary analysis for every variable are presented in Chapter 4.1 and the results of Shapiro-Wilk normality test for each variable are presented in Appendix D.

2.7.2 Kruskal-Wallis one-way analysis of variance

Some data did not follow the normal distribution and were unable to be transformed, such as Metacercariae and Yield in this study. Non-parametric methods are suitable data which is not normally distributed. In this study, the analysis of variance for these data was carried out by the Kruskal-Wallis test, which was widely used in previous research in parasitology and epidemiology. The Kruskal-Wallis test is a non-parametric method for testing whether samples originate from the same distribution (Kruskal and Wallis, 1952; Gregory and Foreman, 2009). It is used for comparing two or more independent samples of equal or different sample sizes. Since it is a non-parametric method, the Kruskal-Wallis test does not assume a normal distribution of the residuals. The null hypothesis is that the medians of all groups are equal, and the alternative hypothesis is that at least one population median of one group is different from the population median of at least one other group. The test statistic is given by:

$$H = (N-1)\frac{\sum_{i=1}^{g} n_i(\bar{r_{i\cdot}} - \bar{r})^2}{\sum_{i=1}^{g}\sum_{j=1}^{n_i}(r_{ij} - \bar{r})^2} \tag{2.7}$$

where: n_i is the number of observations in group i; r_{ij} is the rank of observation j from group i; r_{ij} is the rank of observation j from group i; N is the total number of observations across all groups; $\bar{r_{i\cdot}} = \frac{\sum_{j=1}^{n_i} r_{ij}}{n_i}$ is the average rank of all observations in group i; $\bar{r} = \frac{1}{2}(N+1)$ is the average of all the r_{ij}.

The R 3.5.0 function *kruskal.test*() was applied to compare the variance among groups. The function gives Kruskal-Wallis chi-squared value, Degree of freedom and p-value. At the 95% confidence interval, the critical value of the chi-square statistic with the $df = 4$ degrees of freedom is 9.487729. If the p-value is greater than 0.05 and the chi-squared value is less than 9.487729, the null hypothesis can be accepted, and each group has a statistically equal mean.

The analysis of the Kruskal-Wallis tests results is in Chapter 4.1.

2.7.3 One-way analysis of variance (ANOVA)

For normally distributed data, one-way analysis of variance (ANOVA) was carried out to analyse the difference of variables among fields. The null hypothesis of the ANOVA test is that the fields' means are all equal, while the alternative hypothesis says that at least one pair of the fields' means are all equal, meaning that at least one pair of the fields' means is significantly different. The results of a one-way ANOVA can be considered reliable as long as the following assumptions are met:

- Each set of replicates represents a random sample from different populations.
- Each parent population is normally distributed.
- Each parent population has the same variance.

This required testing the homogeneity of variance before the ANOVA test. The Bartlett's test (Bartlett, 1937) is used to test if samples are from each field with equal variances. If Bartlett's test shows that the data have the same variances (p-value > 0.05), one-way analysis of variance will be conducted. The Bartlett's test is available with the $bartlett.test()$ function and the One-way ANOVA is available with the $aov()$ function in R 3.5.0. The summary of ANOVA results in R function gives F-value, P-value, the degree of freedom and residuals. If the p-value was less than 0.05, we rejected the null hypothesis and can conclude that the fields' means are not all equal and there is at least one pair of fields having a significant difference. The results of the ANOVA are presented in Chapter 4.1.

2.7.4 Pearson's Correlation Coefficients

As part of the exploratory analysis of the relationships between factors in this project, it was of interest to assess the strength of association between the different environmental situations as well as between topographical factors and the number of metacercarial cysts of *F. hepatica*. This was done using Pearson's correlation coefficients. The Pearson's correlation evaluates the linear relationship between two continuous variables. A relationship is linear when a change in one variable is associated with a proportional change in the other variable. The correlation coefficient is given by the formula:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{2.8}$$

The closer r is to +1 or -1, the more closely the two variables are related. If r is close to 0, it means there is no relationship between the variables. If r is positive, it means that as one variable

gets larger the other gets larger. If r is negative it means that as one gets larger, the other gets smaller. The Pearson's Correlation Coefficient was calculated in R using $cor()$ function and choose "$method = pearson$". In this study, we considered thresholds used to interpret r values lists as following:

- $1 < r < 0.5$: A strong positive linear relationship
- $0.5 < r < 0.3$: A moderate positive linear relationship
- $0.3 < r < 0.1$: A low positive linear relationship
- $0.1 < r <$ -0.1: No linear relationship
- -0.1 $< r <$ -0.3: A low negative linear relationship
- -0.3 $< r <$ -0.5: A moderate negative linear relationship
- -0.5 $< r <$ -1: A strong negative linear relationship

2.7.5 Modelling methods to assess relationships between *F. hepatica* and environmental factors

This section discussed the methods used in defining and selection of statistical tests and models along with the process of checking necessary assumptions to assess the relationships between the *F. hepatica* and the environmental factors. All modelling methods and plots of assumption checking were carried out in R.

2.7.5.1 Negative binomial regression model

Regression modeling was used to formally analyse the relationships between *F. hepatica* level and each of the environmental variables in this study. Negative binomial regression models have been widely used in research on the relationship between environment and intensity of *F. hepatica* (Charlier et al., 2014b). For example, Charlier used a negative binomial regression with robust standard deviation to model the small water types and climate data to the *G. truncatula* and *F. hepatica* on four Belgian farms. McCann (McCann et al., 2010b) conducted the univariate models to select suitable variables used in the multiple negative binomial models. A logistic regression model is also widely used in research, especially for modeling environmental factors on *F. hepatica* level (Bennema et al., 2009; Kuerpick et al., 2013; Novobilský et al., 2014).

In this study, the univariate negative binomial regression model for each environmental factor was carried out first and a negative binomial regression model was used to fit all environmental factors to the *F. hepatica* level. Multivariate negative binomial regression models belong to the

family of generalised linear models, except that the response $(\boldsymbol{Y})$ variable is an observed count that follows the negative binomial distribution (McCullagh and Nelder, 1989).The negative binomial regression model is of the following structure:

$$E(\mathbf{Y}_i) = \mu_i = \eta^{-1}(\beta_0 + \beta_i X_i) \quad (2.9)$$

$$\eta = \ln \frac{\mu_i}{1-\mu_i} \quad (2.10)$$

where $E(Y_i)$ is the expected value of Y; $\beta_i X_i$ is a linear combination of unknown parameters X_i; η is the logit link function. In this study, the dependent variable follows the negative binomial distribution, so the logit link function was selected to fit the regression models (Charlier et al., 2014a). In all models, independent variables were used as the nested factors within fields. The Cochno High was used as the reference group and the intercept (β_0) of a model is the mean of the response variable (Y) at the mean of the independent variable (X_i)for the reference group.

The R function *glm()* was used to fit the negative binomial generalised linear regression model. The dependent and independent variables used in fitting models was listed as follow:

- Response variable: Yield2 (Metacercariae per 0.1 gram)

- Independent variables:
  Dataset 2: Remote sensing indices:
  Normalized difference water index (NDWI) and Normalized difference vegetation index (NDVI)
  Dataset 3: Topographical factors:
  Elevation, Slope, and Aspect
  Dataset 4: Soil properties:
  Soil Moisture and Soil Temperature
  Dataset 5: Pasture grass:
  Grass Height and Grass Weight

2.7.5.2 McFadden's Pseudo R^2

When fitting the logistic regression model, R^2-type measures often appear very small even when other measures suggest that the model fits the data well. However, to evaluate the goodness-of-fit of logistic models, several pseudo R^2s have been developed. These are 'pseudo' R^2s because they look like R^2 in the sense that they are on a similar scale, ranging from 0 to 1 with higher values indicating better model fit. This study used the McFadden's Pseudo R^2. McFadden's Pseudo R^2 measure (McFadden, 1973) is defined as

$$R^2_{McFadden} = 1 - \frac{(\log L_c)}{\log(L_{null})} \tag{2.11}$$

where L_c denotes the (maximised) likelihood value from the current fitted model, and L_{null} denotes the corresponding value but for the null model, the model with only an intercept and no covariates. If comparing two models on the same data, McFadden's would be higher for the model with the greater likelihood. The $PseudoR2()$ function in R 3.5.0 package BaylorEdPsych are used to calculate the McFadden's Pseudo R^2.

2.7.5.3 Akaike information criterion (AIC)

The Akaike information criterion (AIC) (Akaike, 1974, 1987) is an estimator of the relative quality of statistical models for a given set of data. Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. Thus, AIC provides a means for model selection. The AIC was widely used to evaluate the goodness of fit of a regression model, the preferred model usually having the minimum AIC value. The definition of AIC is

$$AIC = 2k - 2\ln\left(\hat{L}\right) \tag{2.12}$$

Where k is the number of estimated parameters in the model and L is the maximum value of the likelihood function for the model.

2.7.5.4 Chi-Squared goodness of fit test

Chi-Squared goodness of fit test (Pearson, 1992) is a statistical test applied to sets of categorical data to evaluate how likely it is that any observed difference between the sets arose by chance. It is suitable for unpaired data from large samples. It is the most widely used of many chi-squared tests. Research shows that chi-squared test fits negative binomial distribution data better (Charlier et al., 2014b). For a chi-square goodness of fit test, the hypotheses take the following form:

- H_0: there is no significant difference between the observed and the expected value
- H_a: there is a significant difference between the observed and the expected value

A chi-square random variable (χ^2) defined by the following equation

$$\chi^2 = \sum[(O_i - E_i)^2/E_i] \tag{2.13}$$

where O_i is the observed value and E_i is the expected value. In this study, the $anova()$ function in R 3.5.0 package MASS was used to measure the goodness of each independent variables to the model and select *test='Chisq'*. The result of chi-squared goodness of fit test for each variable in

the model was given, including the chi-squared valued, degrees of freedom, deviance residual, degree of freedom for residuals and p-value. At the 95% confidence interval, if the p-value is less than 0.05, the null hypothesis will be rejected.

Chapter 3

Preliminary Risk Maps

3.1 Risk classification

In this study, normalized difference water index, normalized difference vegetation index, slope, and soil parameters were used to evaluate the risk of occurrence *F. hepatica*. For the soil layer, the criteria were soil texture and soil particle size. These soil properties can determine the drainage ability of an area. Soil with fine particles often provides poorer drainage than soil with coarser particles (McCann et al., 2010a). Hence, areas with smaller particle sizes were identified as high-risk areas. Based on this, the soil texture classified in seven class from high risk (7) to low risk (1), showing in Table 3.2. In the Table 3.2, the soil texture with ND means no data. NDVI and NDWI indicate the vegetation and the land surface moisture. The higher NDVI values equate to higher green plant coverage. High NDWI values indicate that the surface moisture of the land is high (Bennema et al. (2009); Tum et al. (2004)). Above all, the high NDVI and NDWI areas were identified as the high-risk areas. The slope also influences the land surface drainage system. The flat areas represent poor at drainage, especially in a wet climate like Scotland. This means areas with less slope were classified as potential risk areas of *F. hepatica*. Hence, the NDWI, NDVI and slope were classified in the risk class from high (5) to low (1) in Table 3.2.

Figure 3.1 presents the classification maps of NDVI, NDWI, Slope and Soil texture at Cochno. Figure 3.2 shows the classification maps at Dumgoyne. For each field, add each risk map together and calculate a total risk level. Then reclassify the total risk level from high to low to generate a preliminary risk map for each field. Figure 3.3 shows the preliminary risk maps of both Cochno and Dumgoyne.

Table 3.1: Risk class for the data layers (Risk decreasing from 5 to 1)

Risk Class	NDVI	NDWI	Slope (degree)
5	0.717-0.858	0.44-0.553	0-5
4	0.576-0.717	0.328-0.44	5-10
3	0.435-0.576	0.216-0.328	10-15
2	0.294-0.435	0.104-0.216	15-30
1	<0.294	<0.104	>30

Table 3.2: Soil texture of study areas and the risk class (Risk decreasing from 7 to 1)

Soil_TEXT	EST_TEXT	Grain-size Name	Particle Diameter(mm)	Risk Class
S_NL	Sand>Loam	Fine	<0.25	7
S_SL	Sand	Fine	<0.25	6
L_C_S	Loam>Clay>Sand	Argillic-Arenaceous	<2.0	5
NL	Loam>Clay>Sand	Medium	0.25>2	4
C_S	Clay>Sand	Mixed (Argillic-Arenaceous)	ALL	3
S_L	Sand>Loam	Mixed (Argillic-Arenaceous)	ALL	2
ALL	All	Peat	Not applicable	1
ND	–	–	–	No Data

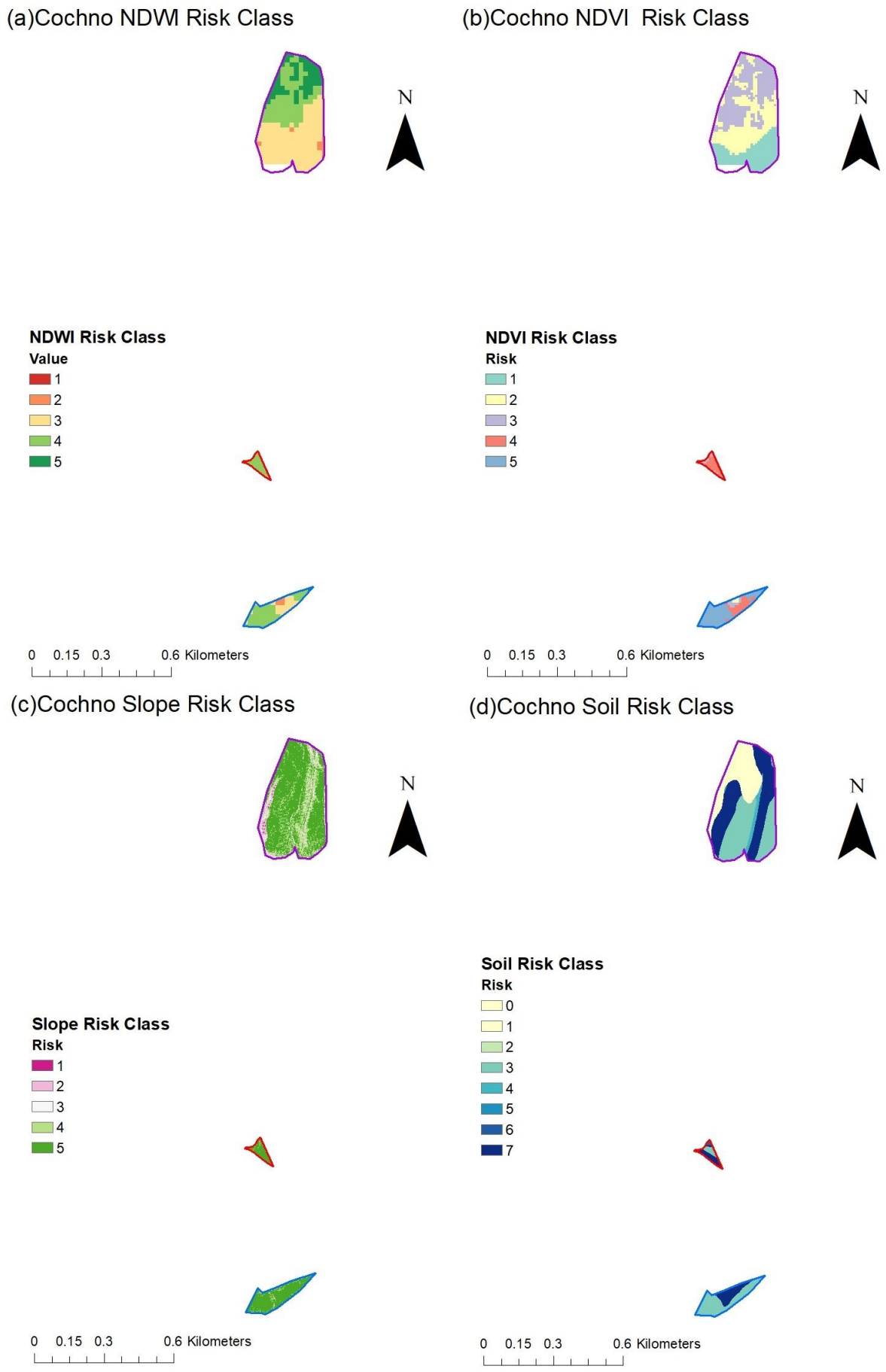


Figure 3.1: Risk Classification Maps of Cochno: (a) NDWI; (b) NDVI; (c) Slope; (d) Soil.

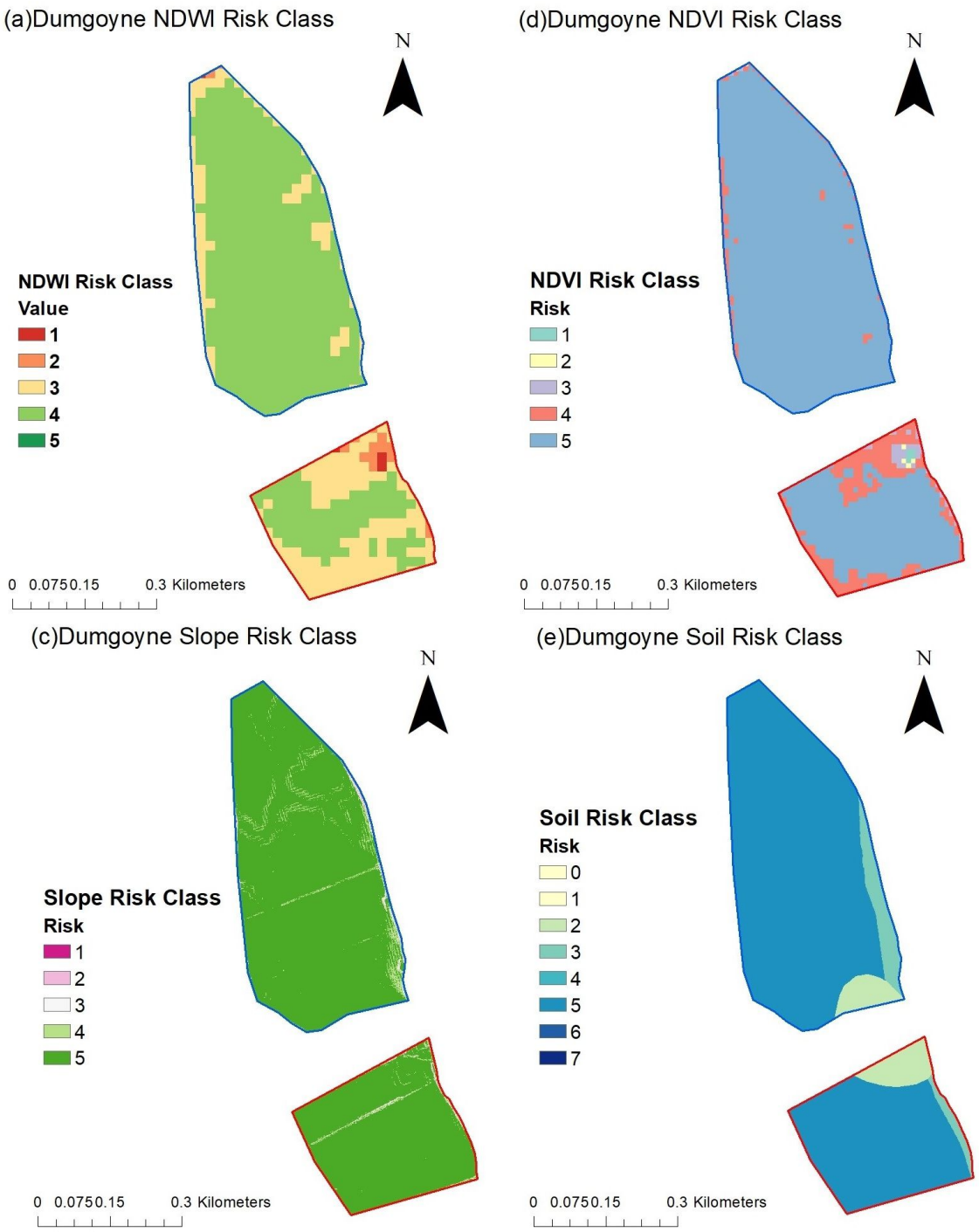


Figure 3.2: Risk Classification Maps of Dumgoyne: (a) NDWI; (b) NDVI; (c) Slope; (d) Soil.

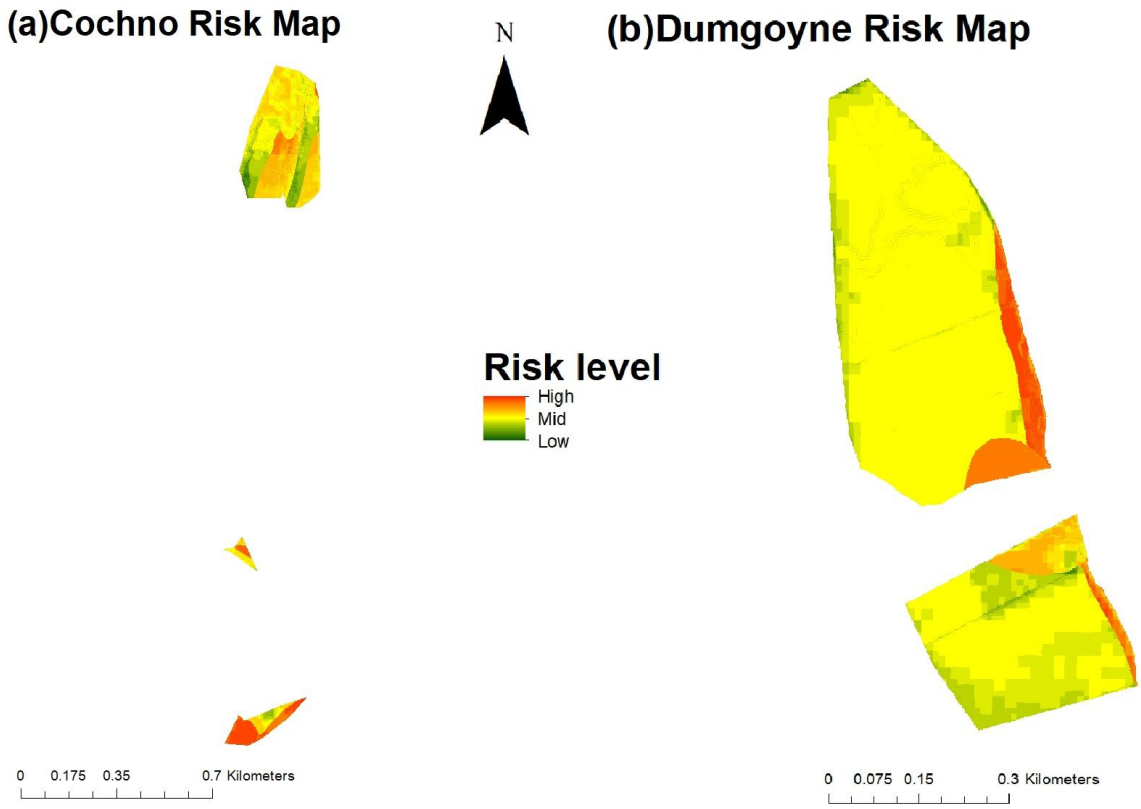


Figure 3.3: Risk Maps: (a) Cochno; (b) Dumyoyne.

3.2 Sampling locations

The preliminary risk map was used as one of the references on the selection of sampling locations. To verify the accuracy of the risk map, the sampling locations were selected not only from high-risk and mid-risk areas but also on low-risk areas. the counts of metacercarial cysts from each sampling point were obtained in the laboratory analysis. Figure 3.4 (a) to (c) shows the preliminary risk map and the metacercarial cyst counts for the sampling points in Cochno. It was shown that some areas in Cochno which were assigned a preliminary high risk status did not have many metacercaria and some areas predicted to be low risk yielded a very high number of metacercariae. For example, in Cochno Low, L01 was located in the mid-risk area; how- ever, the number of the metacercariae found at the L01 was 71.43. This is a relatively high level of metacercariae. The same situation occurred in L06 as well. In addition, some sampling locations in the high-risk areas returned low metacercarial cyst counts, such as L02 and L08. Another two Cochno fields were the same. H01 in Cochno High and M05 in Cochno Mid were misclassified into the low-risk areas. Figure 3.4 (d) and (e) shows the relationship between the preliminary risk map and the metacercarial cyst counts in Dumgoyne. The risk map identified that most of the areas in Dumgoyne were mid or low risk. For example, in Dumgoyne North, N11 was located in a mid-risk area, while the number of metacercarial cysts was extremely high

(around 386). N07 was the opposite situation, which was classified into very high-risk areas but decidedly fewer metacercarial cysts were found.

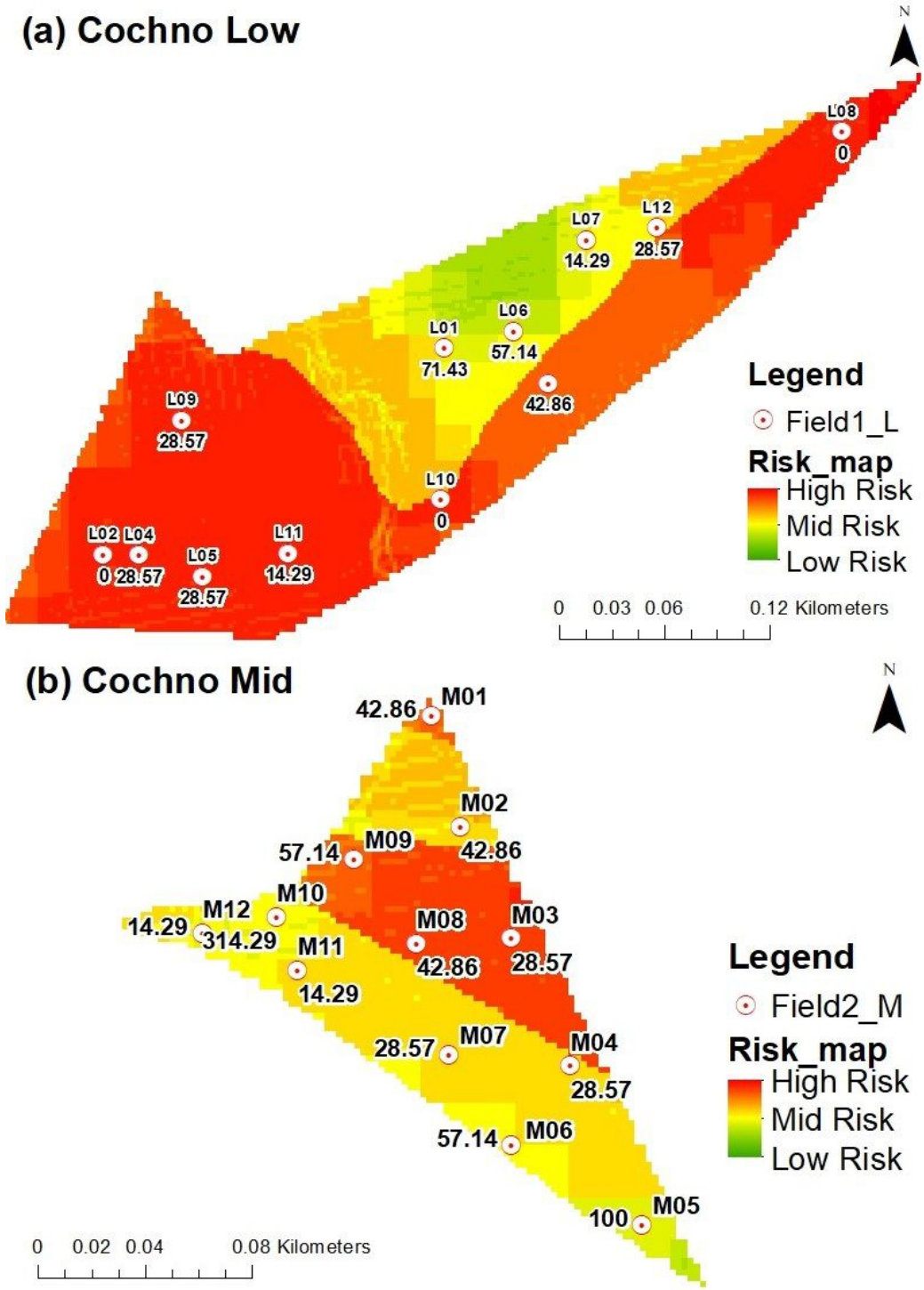
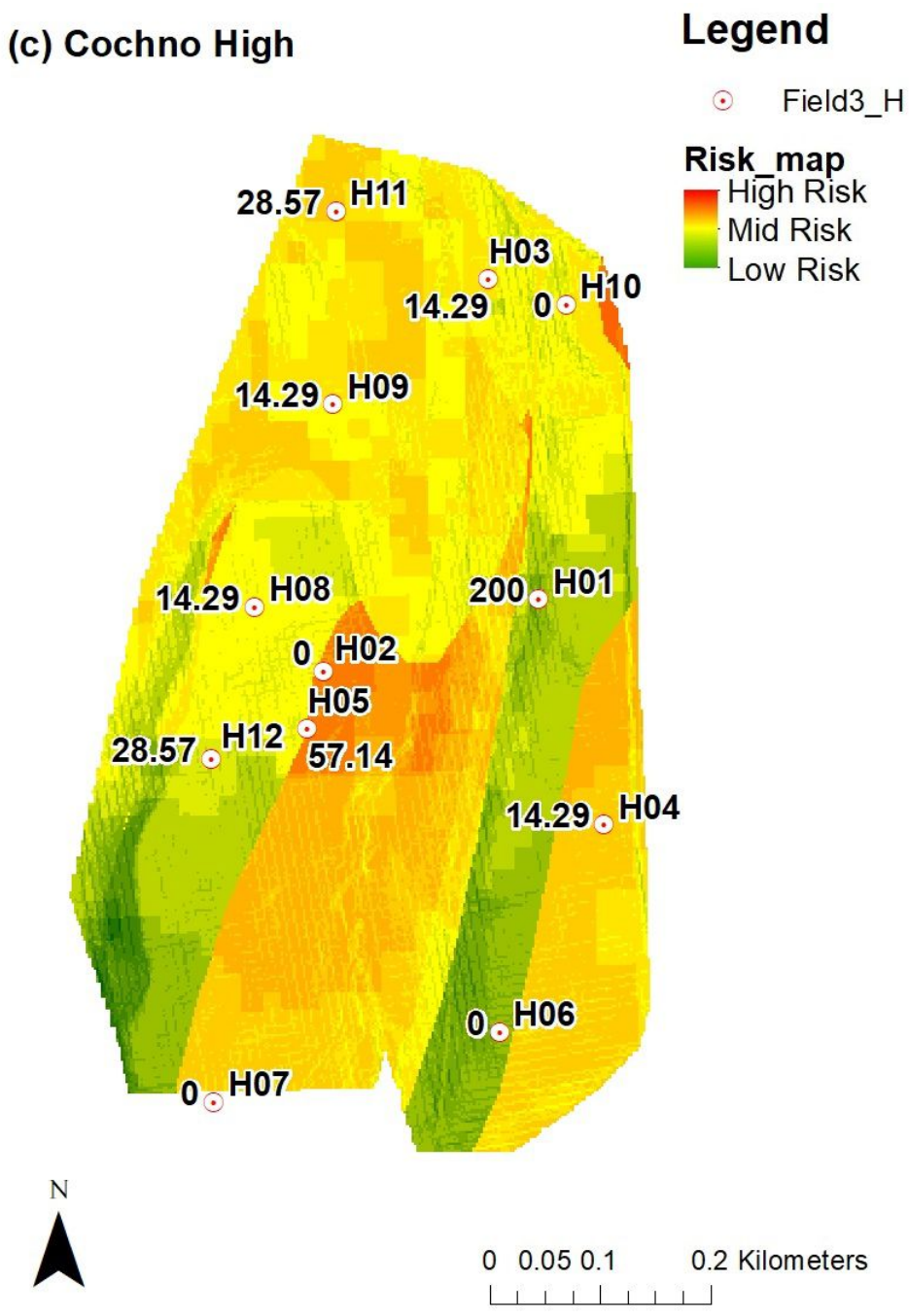
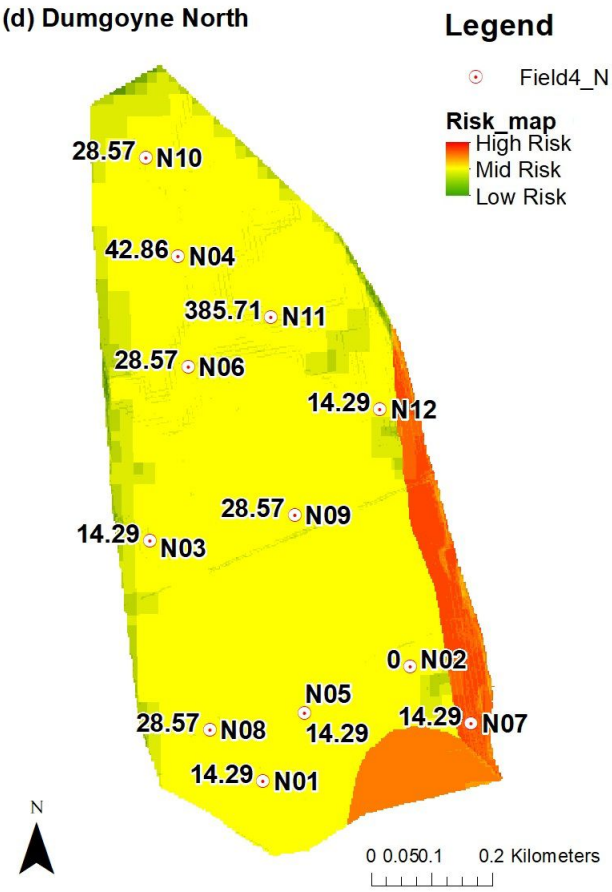
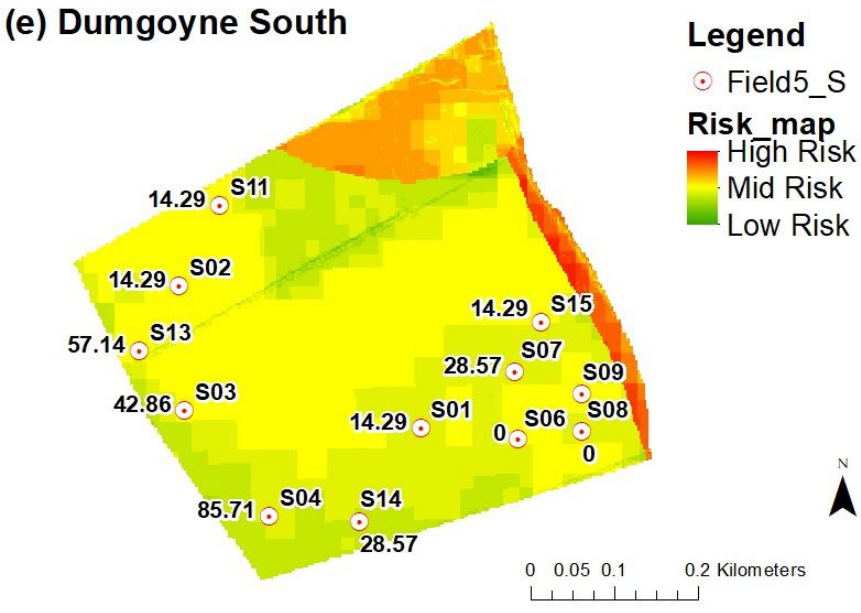


Figure 3.4: Risk maps and metacercarial cyst counts: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South.


(c) Cochno High
Legend
Field3_H
Risk_map
High Risk
Mid Risk
Low Risk
28.57 H11
H03
14.29
0 H10
14.29 H09
14.29 H08
200 H01
0 H02
H05
28.57 H12
57.14
14.29 H04
0 H06
0 H07
N
0 0.05 0.1 0.2 Kilometers


(d) Dumgoyne North
Legend
Field4_N
Risk_map
High Risk
Mid Risk
Low Risk
28.57 N10
42.86 N04
385.71 N11
28.57 N06
14.29 N12
28.57 N09
14.29 N03
0 N02
N05
28.57 N08
14.29
14.29 N07
14.29 N01
N
0 0.050.1 0.2 Kilometers


(e) Dumgoyne South
Legend
Field5_S
Risk_map
High Risk
Mid Risk
Low Risk
14.29 S11
14.29 S02
14.29 S15
57.14 S13
28.57 S07
S09
42.86 S03
14.29 S01
S06
S08
0
0
85.71 S04
S14
28.57
N
0 0.05 0.1 0.2 Kilometers

Chapter 4

Exploratory Analysis of Environmental Factors and Density of *F. hepatica* in pasture

4.1 Initial exploration of *F. hepatica* and environmental factors among fields

This section presents the results of the statistical analysis of all variables and an analysis of the variation of each variable among the five study areas. There are five categories including 11 variables and 58 samples within the five study areas which are used in the analysis.

4.1.1 *F. hepatica* density in pasture

4.1.1.1 Metacercarial cyst counts

Metacercarial cyst counts per 100 ml solution can represent the *F. hepatica* density within each study areas. Figure 4.1 shows the metacercarial cyst counts in each of sampling locations around the fields. The highest metacercarial cyst counts was 386 in N11, Dumgoyne North, and the lowest metacercarial cysts was zero. The mean count of metacercarial cyst counts in Cochno Mid were higher than other areas.

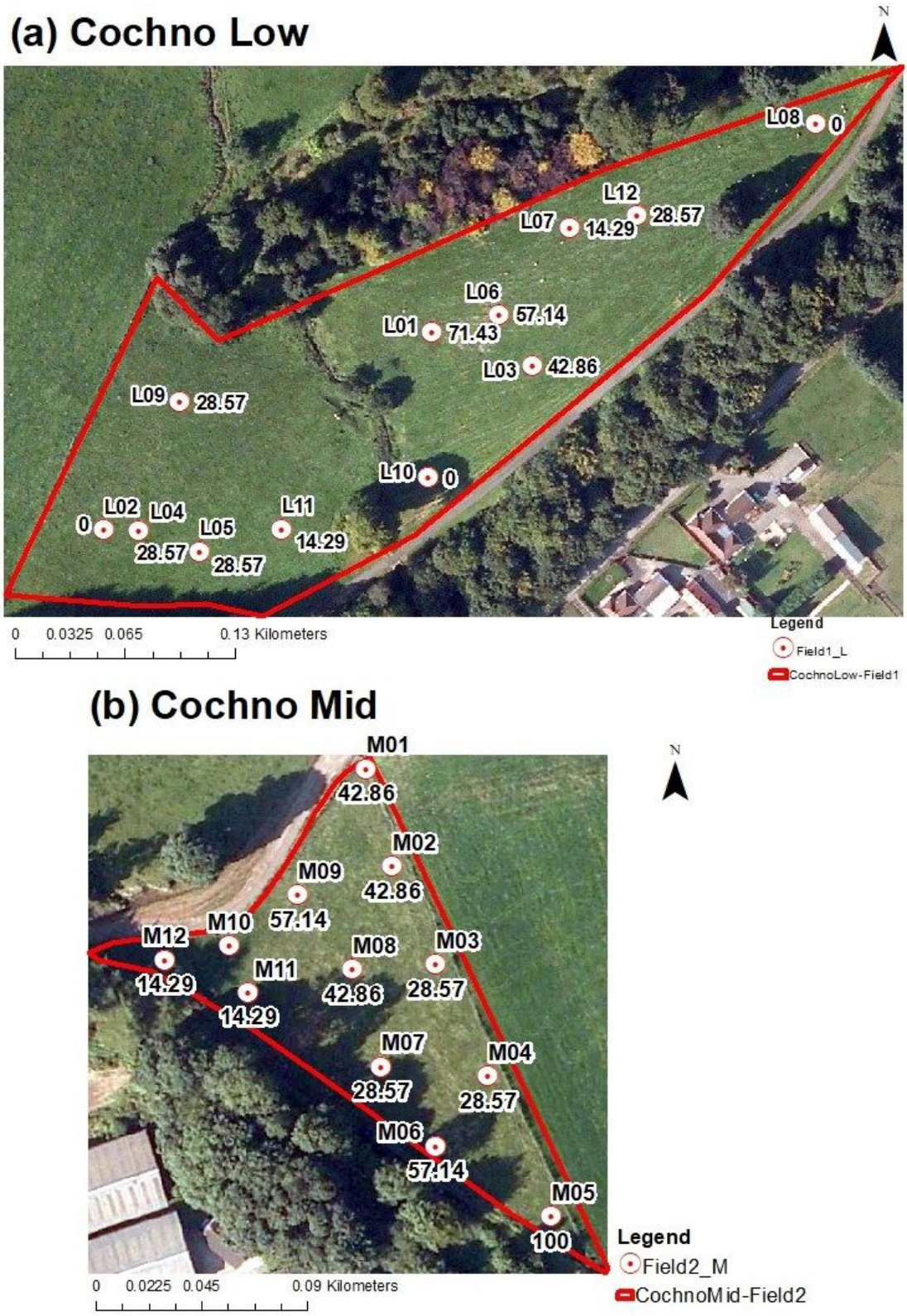
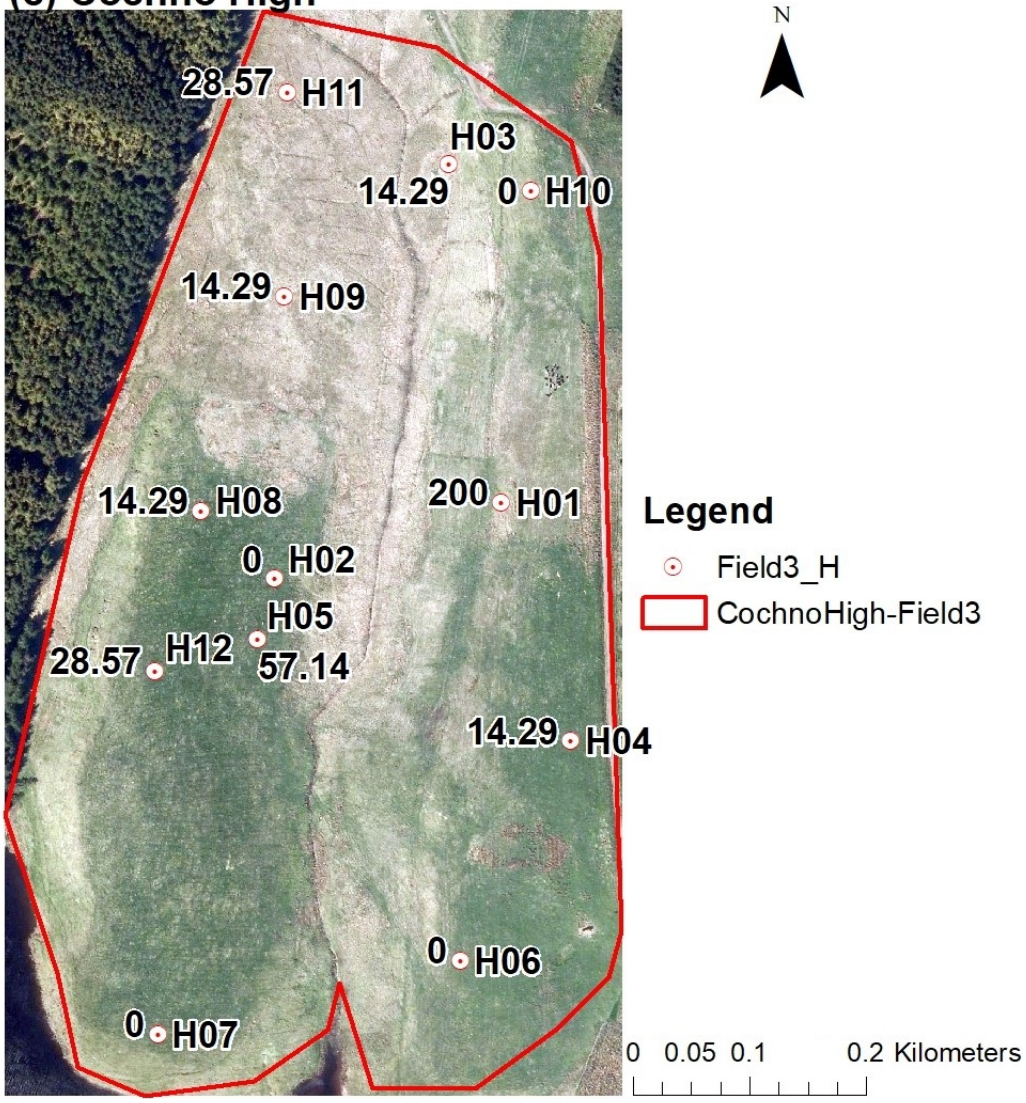
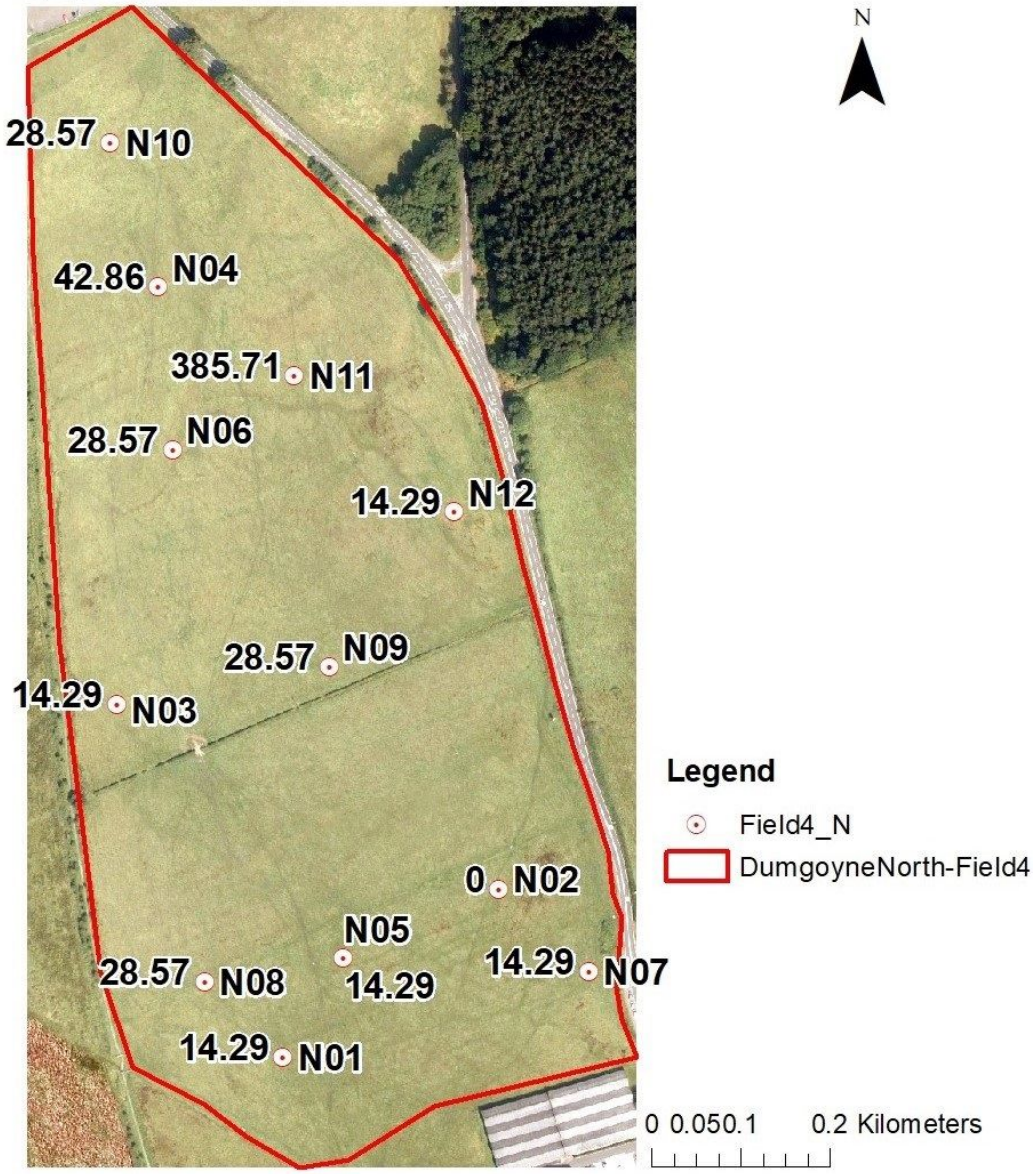


Figure 4.1: The metacercarial cyst count maps for sampling points within the study fields: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South.

(c) Cochno High



(d) Dumgoyne South



(e) Dumgoyne South

The histogram of metacercarial cysts (Figure 4.2) for all 59 samples shows a negative binomial distribution, as expected. The number of metacercarial cysts between 0 to 100 had the highest frequency and lower frequency was between 300 to 400 (Figure 4.2 (f)). Although the metacercarial cysts followed a similar distribution pattern within fields, except Cochno Low (Figure 4.2 (a)), the range of metacercarial cysts varied across fields.

Each field's metacercarial cyst profile can be seen in the kernel density curve in Figure 4.3. The red line represents the mean count of metacercariae of all samples and the blue line is the median of all samples. A Kruskal-Wallis rank sum test was carried out to determine the effect of Field on metacercarial cysts count. The value of the test statistic was 7.8137. The p-value (0.099) was greater than 0.05, under 95% confidence interval; also, the value of the test statistic was less than the chi-square-tabulation (9.49). The conclusion is, therefore, that the null hypothesis H0 was accepted: the means of the 5 groups were statistically equal. Hence, there was no significant difference among the five fields in the number of metacercarial cysts.

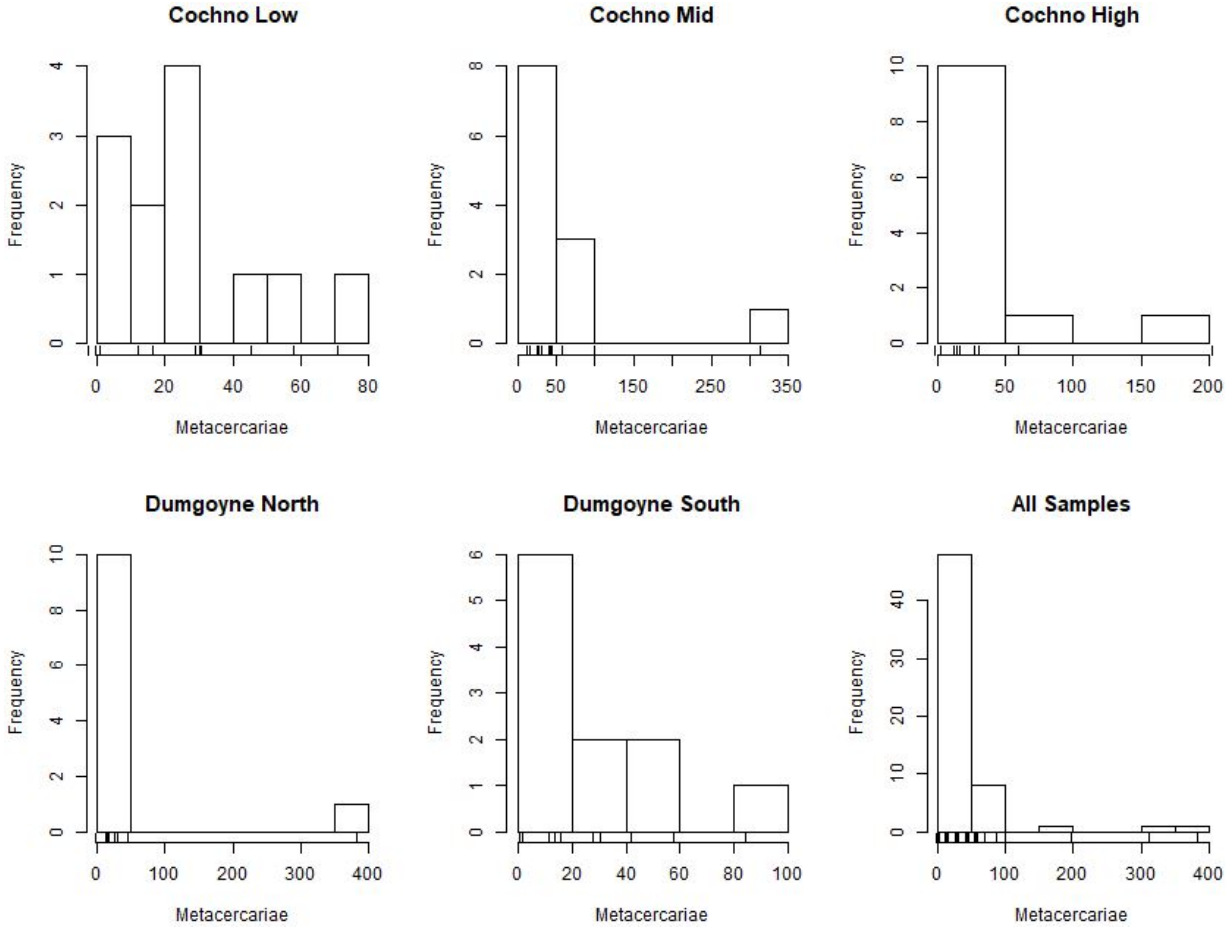


Figure 4.2: Histograms of metacercarial cyst for each field and all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples.

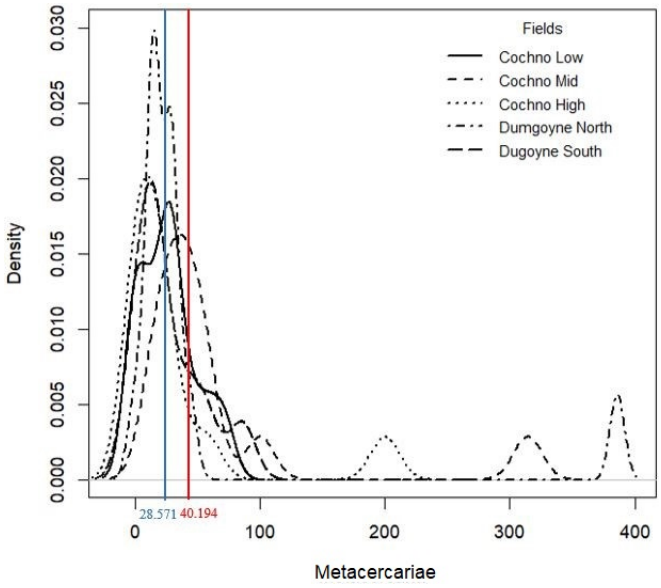


Figure 4.3: Density curve of metacercarial cysts for each field. Red vertical line is the overall average count. Blue vertical line is the median.

4.1.1.2 Yield (Metacercariae/g)

Yield is used to mean the metacercariae per gram pasture in our study, which reflects the contamination of forage grass by metacercariae. The Yield map of each field is presented in Figure 4.4. Most Yield values were under 1. The highest yield value was 1.31, the M10 in Cochno Mid, also the only yield value greater than 1, while the highest metacercarial cyst was in Dumgoyne North.

As shown in Figure 4.5 (f), the yield followed the negative binomial distribution and Figure 4.5 (a) to (e) displays the histogram of each field showing the same distribution pattern with total samples but also representing the variation among fields. The variation of yield values among the fields is presented in the kernel density curve (Figure 4.6). The red line is the mean yield value of all samples, and the blue line is the median of all samples. It is shown that the density curve of yield for each field is highly skewed. There were differences in the probability of data distribution among fields. For example, the yield in Dumgoyne North had a very high density , in the range of 0 to 0.05 , and another high density occurred in the range of 0.4 to 0.45. A Kruskal-Wallis rank sum test was carried out to determine the variation of yield (metacercariae/g) among fields. The value of the test statistic was 11.286. The p-value (0.024) was less than 0.05; also, the value of the test statistic was greater than the chi-square-tabulation. Therefore, the means of the five groups were not statistically equal, and yield values among the five fields were significantly different.

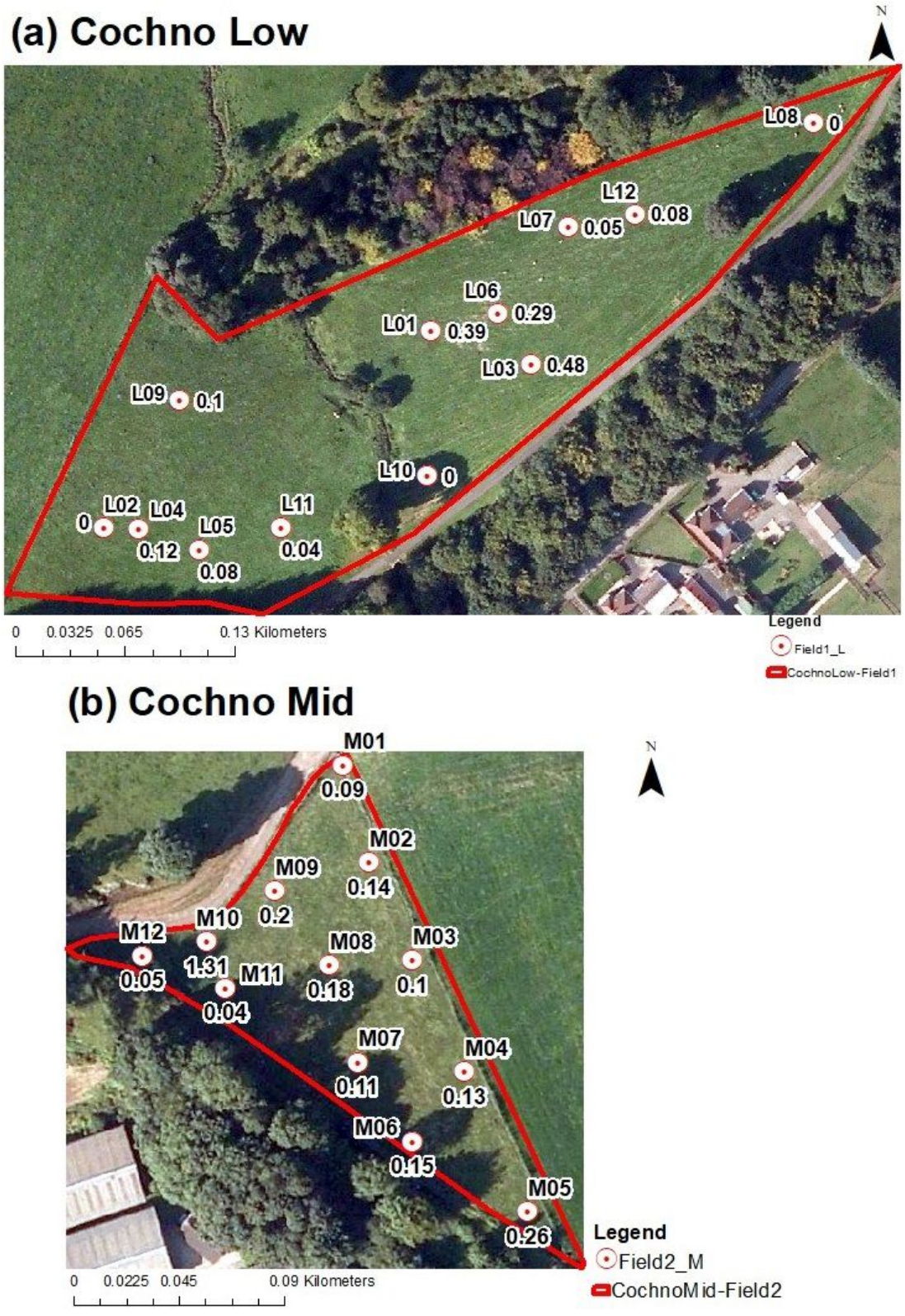
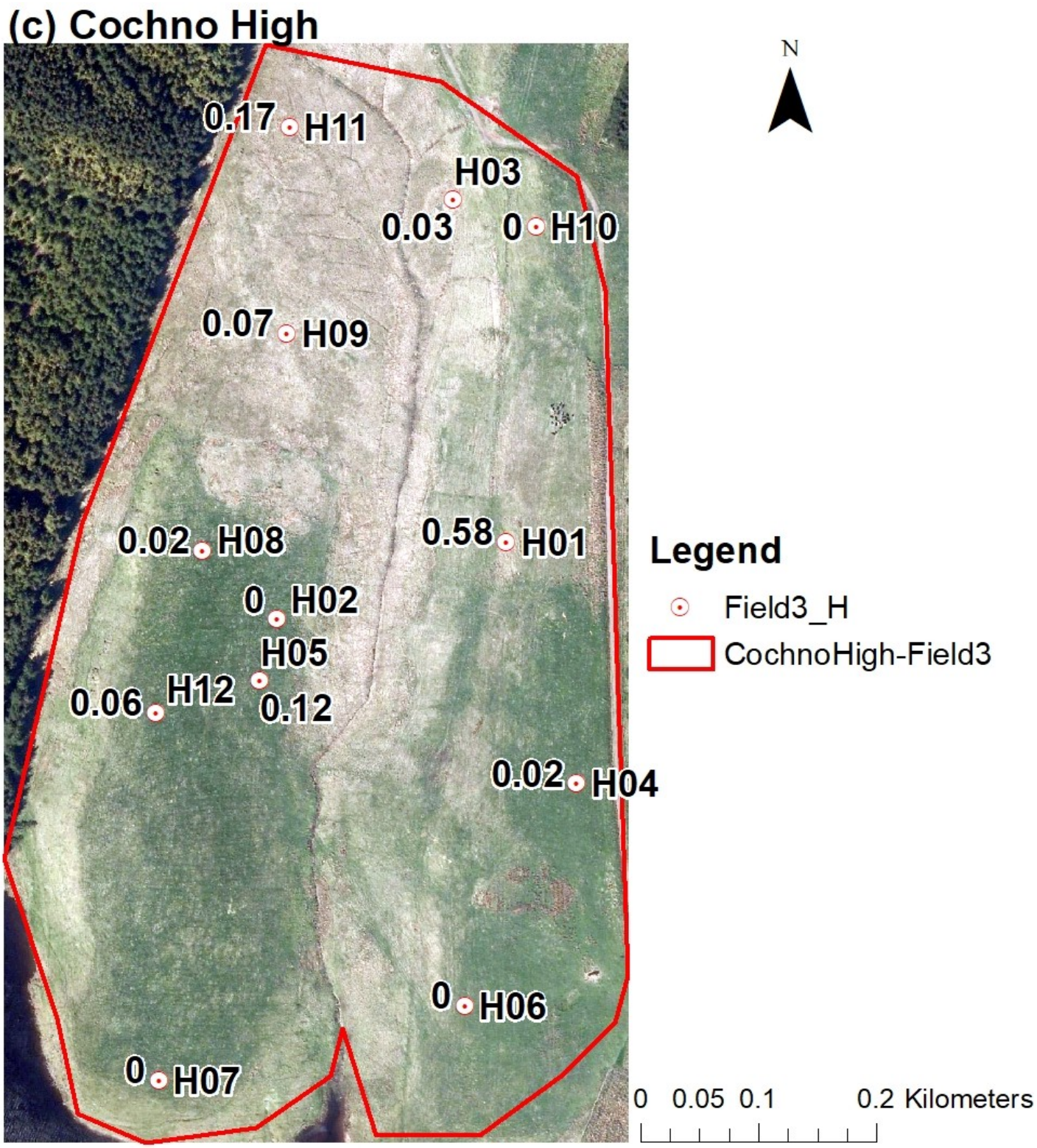
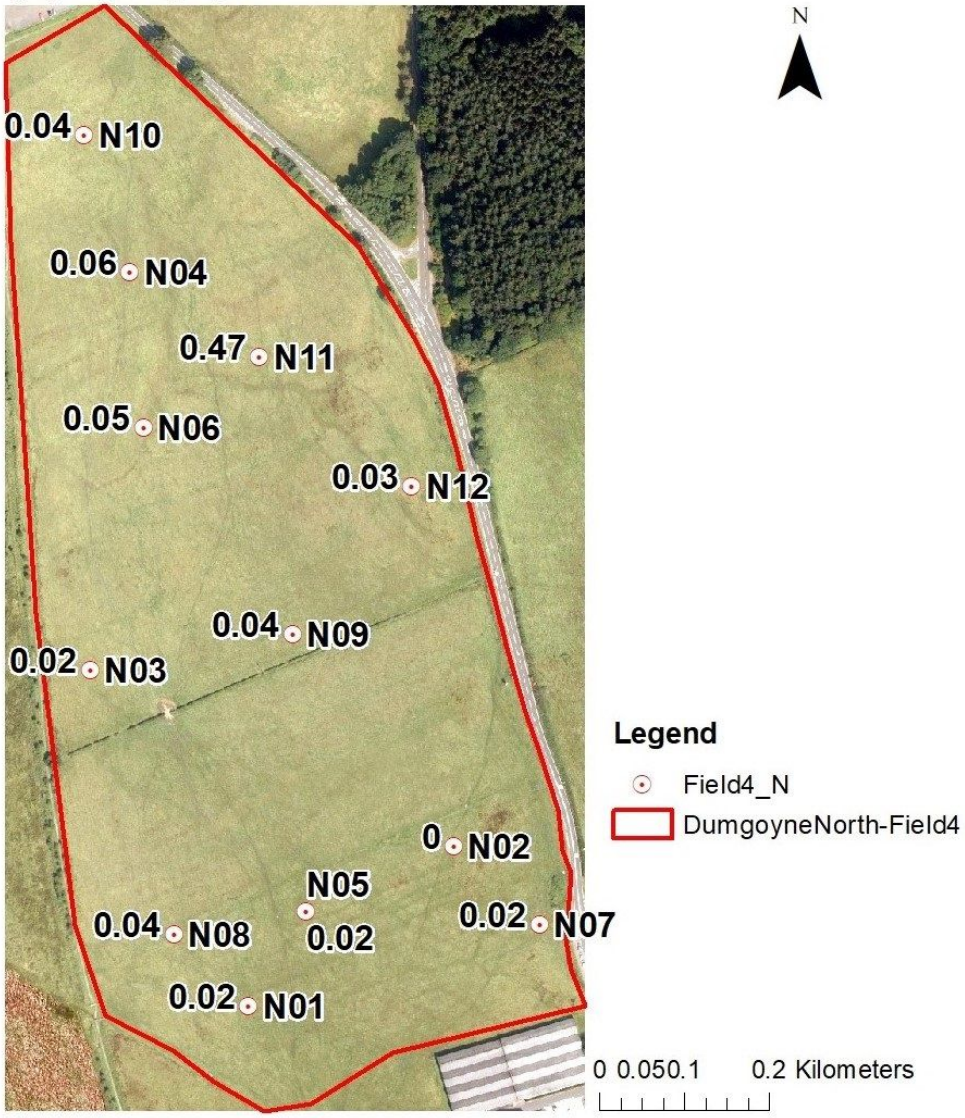


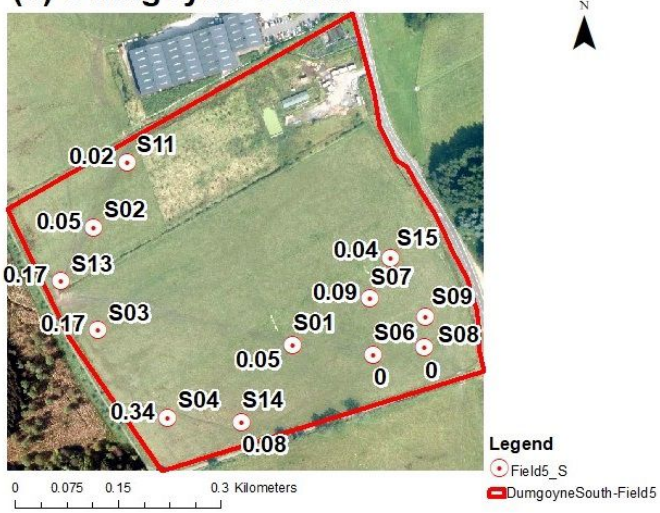
Figure 4.4: The yield for sampling points within the study fields: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South.


(c) Cochno High
N
0.17
H11
H03
0.03
0
H10
0.07
H09
0.02
H08
0.58
H01
0
H02
H05
0.06
H12
0.12
0.02
H04
0
H06
0
H07
Legend
Field3_H
CochnoHigh-Field3
0 0.05 0.1 0.2 Kilometers

(d) Dumgoyne South



(e) Dumgoyne South



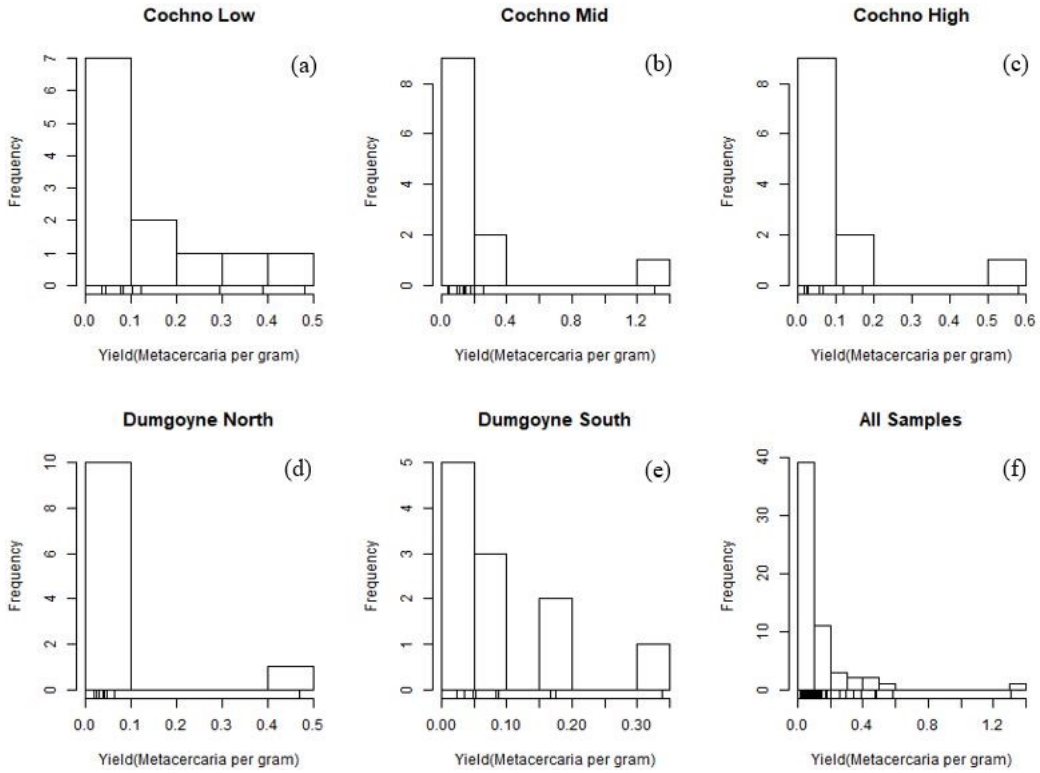


Figure 4.5: Histograms of yield for each field and for all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples.

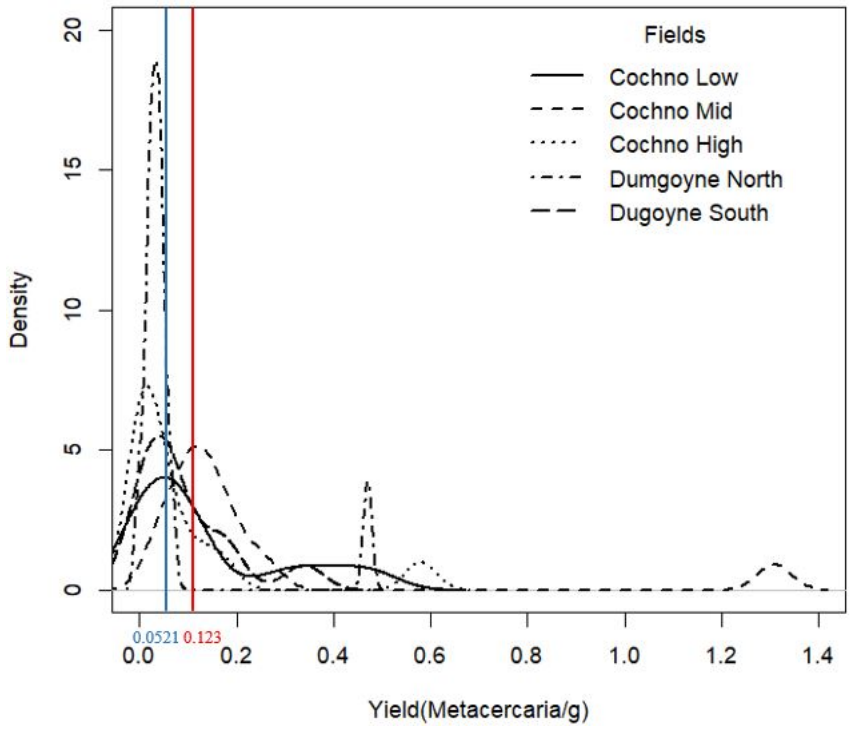


Figure 4.6: Density curve of yield (metacercariae/g) for each field.

4.1.2 Remote sensing indices

4.1.2.1 Normalized difference water index

Normalized difference water index (NDWI) indicates land surface moisture. The NDWI values for 31August 2017 were used in the analysis and displayed in Figure 4.7 with the NDWI value for each sampling point labelled. The highest NDWI of all samples was 0.46 in Cochno High (H11), and the lowest was 0.201 in the Cochno Low (L06). According to the Shapiro-Wilk normality test result, NDWI values did not follow the normal distribution, and the histogram of NDWI in Figure 4.8 supported some of the features of the data distributions. There were variations among the distribution pattern of fields in Figure 4.8. All NDWI values were positive, and the most of the NDWI values were in the range from 0.35 to 0.45. Figure 4.9 shows that the density curves of the NDWI in each field were highly skewed to the right, except for the Cochno Mid field. Consequently, a Box-cox Transformation was carried out to transform NDWI and reduce the skewness of data. From the Figure 4.6, the Dumgoyne North had the narrowest NDWI range, and Cochno Low had a similar curve to Cochno High. The red line is the mean of all samples, and the blue line is the median of all samples. The Kruskal-Wallis test was conducted to determine the difference among fields. The p-value given by test (0.039) was lower than 0.05, and we can conclude that the NDWI differed significantly within the fields.

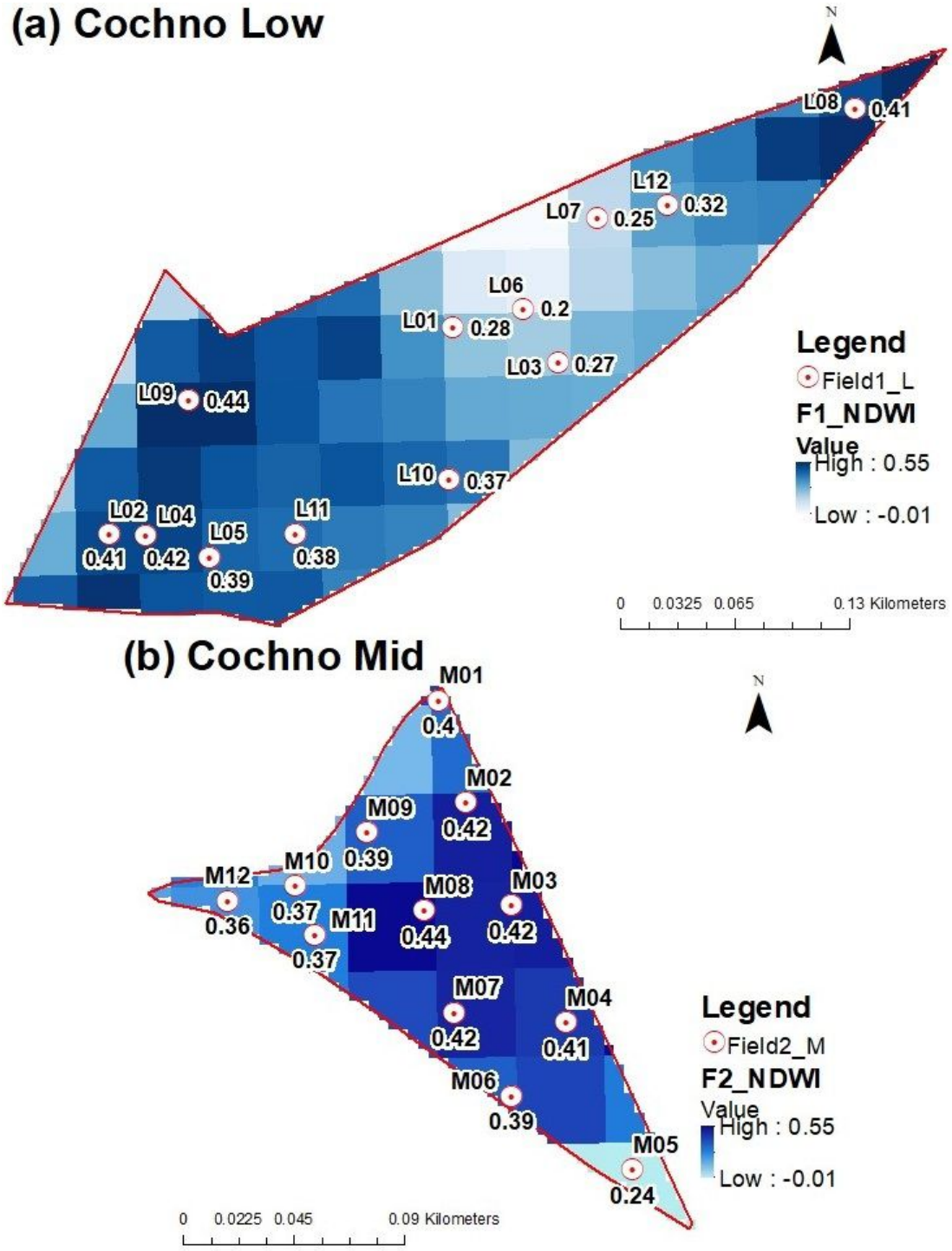
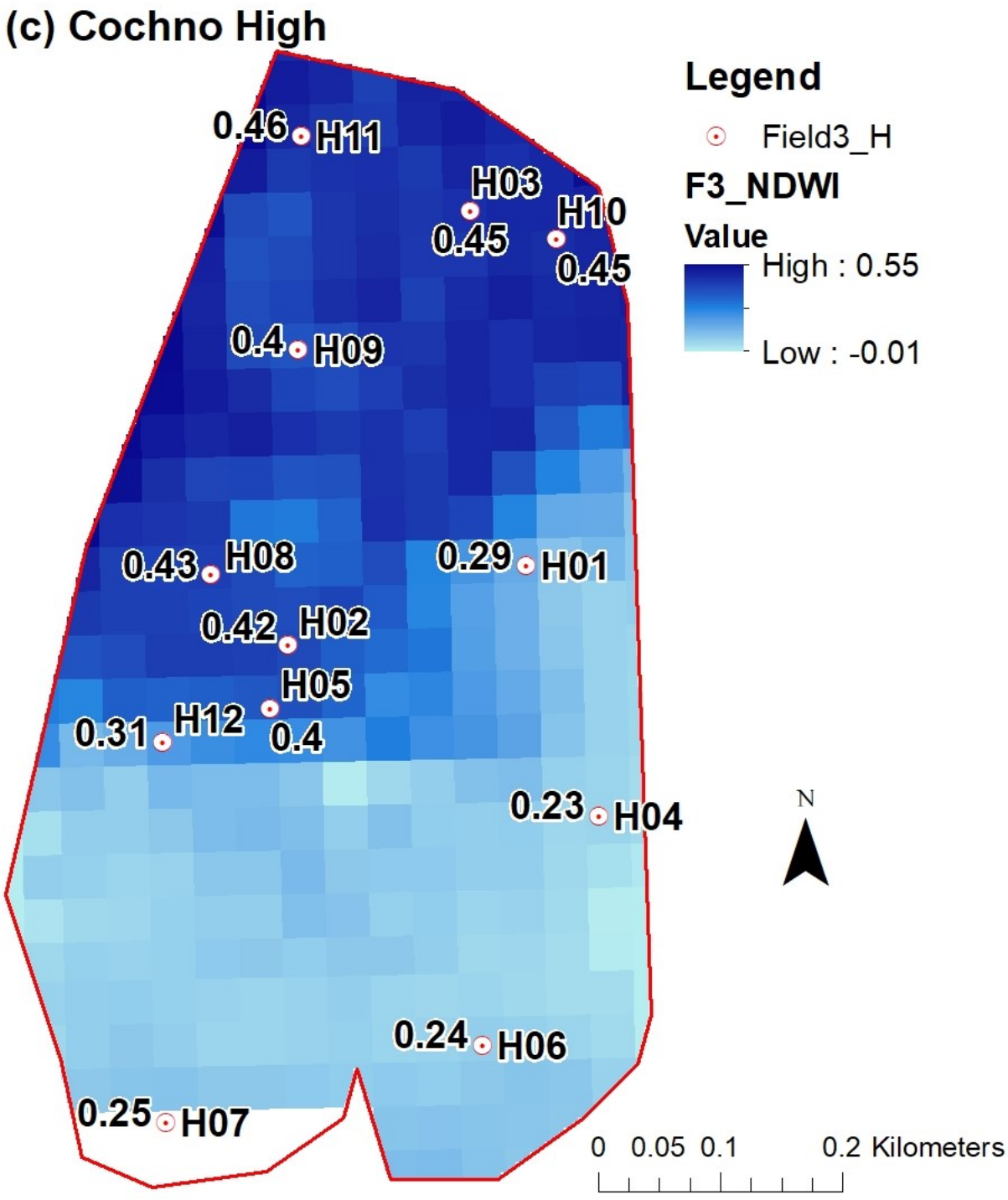
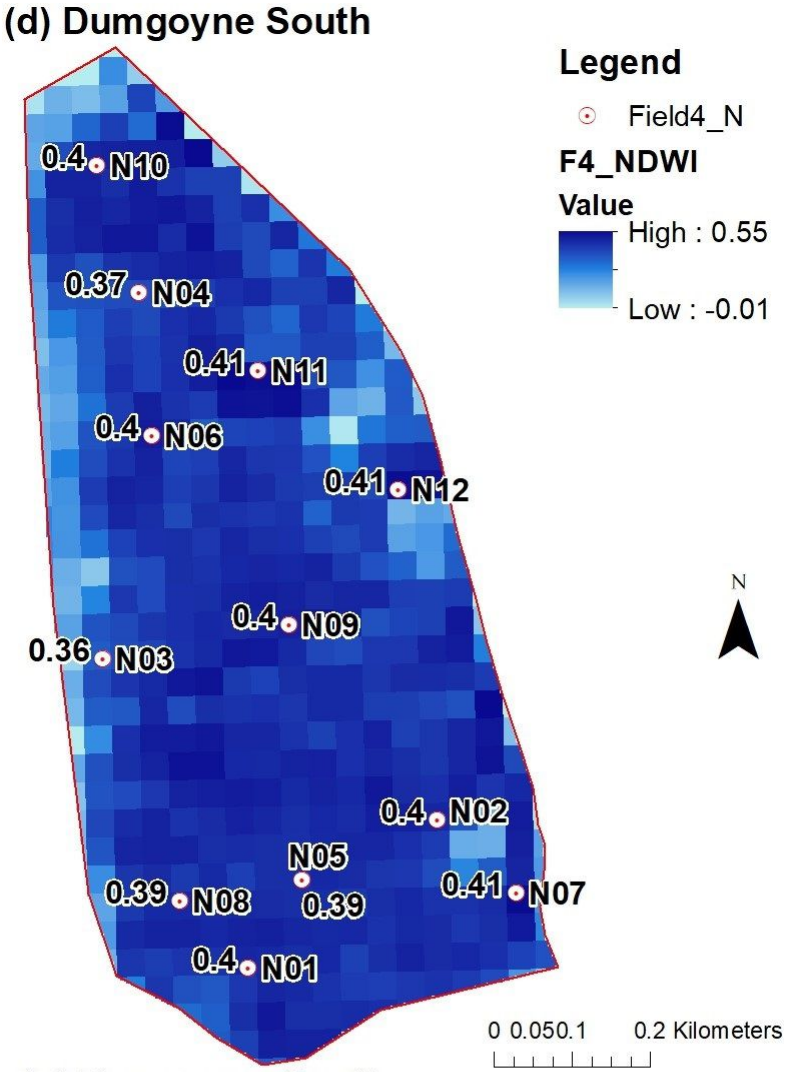
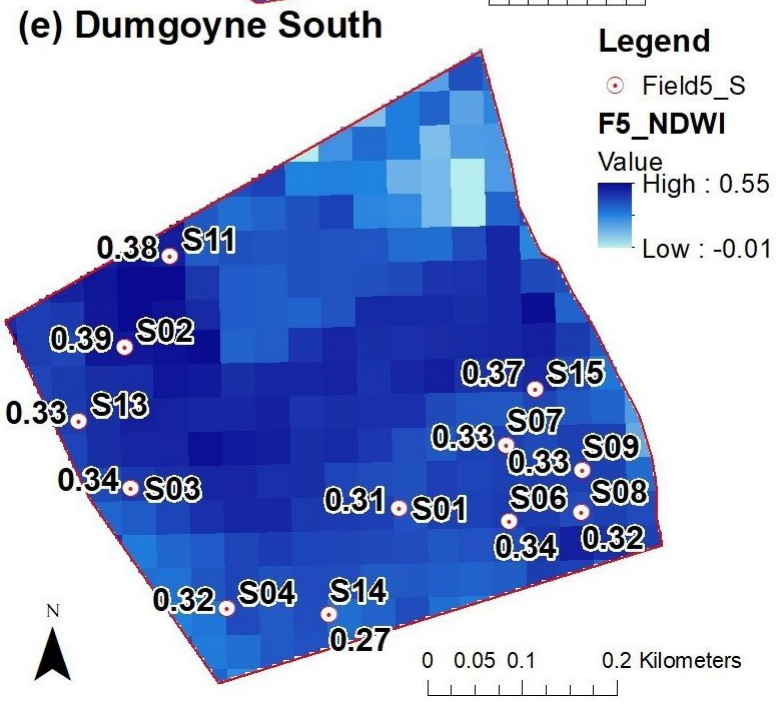


Figure 4.7: The NDWI maps for study fields and the NDWI of each sampling point: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South.


(c) Cochno High
Legend
Field3_H
F3_NDWI
Value
High : 0.55
Low : -0.01
0.46 H11
H03
0.45
H10
0.45
0.4 H09
0.43 H08
0.29 H01
0.42 H02
H05
0.31 H12
0.4
0.23 H04
0.24 H06
0.25 H07
N
0 0.05 0.1 0.2 Kilometers


(d) Dumgoyne South
Legend
Field4_N
F4_NDWI
Value
High : 0.55
Low : -0.01
0.4 N10
0.37 N04
0.41 N11
0.4 N06
0.41 N12
0.4 N09
0.36 N03
0.4 N02
N05
0.39 N08
0.39
0.41 N07
0.4 N01
N
0 0.050.1 0.2 Kilometers


(e) Dumgoyne South
Legend
Field5_S
F5_NDWI
Value
High : 0.55
Low : -0.01
0.38 S11
0.39 S02
0.37 S15
0.33 S13
S07
0.33 S09
0.33
0.34 S03
0.31 S01
S06 S08
0.34
0.32
0.32 S04
S14
0.27
N
0 0.05 0.1 0.2 Kilometers

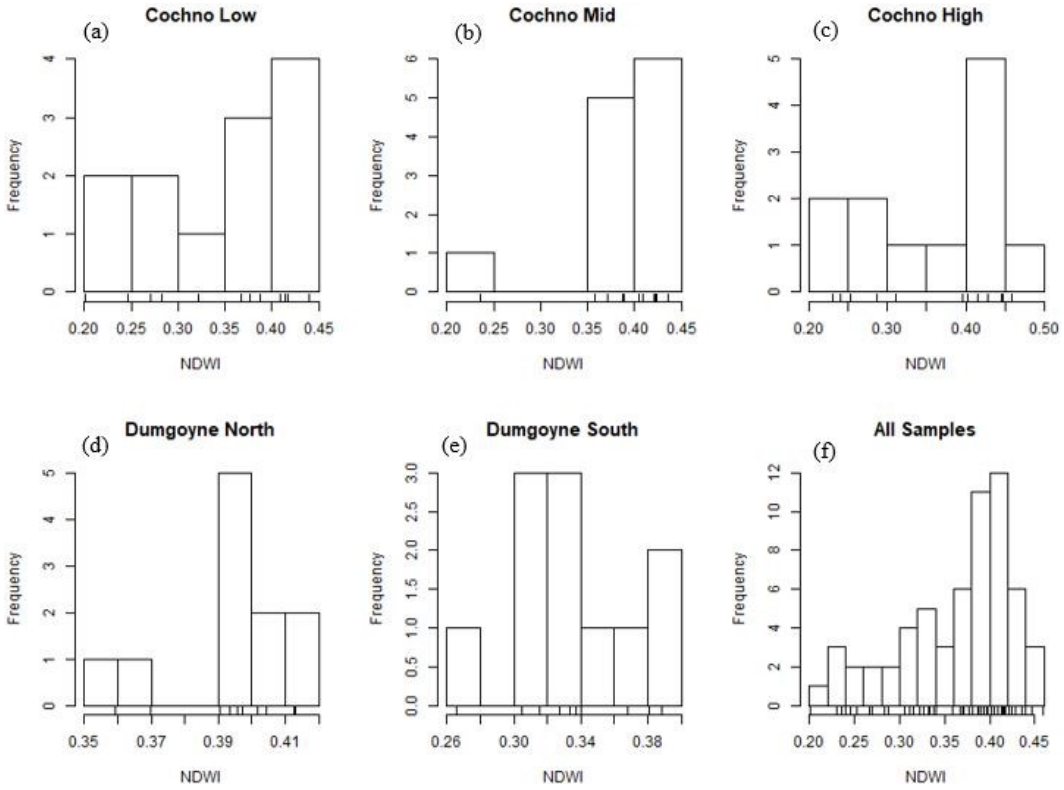


Figure 4.8: Histograms of NDWI for each field and all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples.

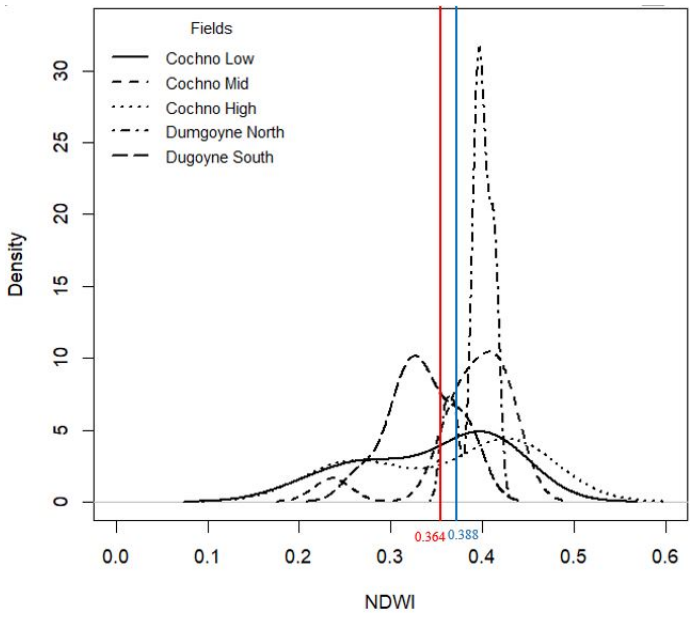


Figure 4.9: Density curve of NDWI for each field.

4.1.2.2 Normalized difference vegetation index

The Normalized difference vegetation index (NDVI) reflects the growth and abundance of green vegetation. In this study, the NDVI values were acquired from the Sentinel-2 satellite image of 31 August 2017. There were 60 NDVI samples in total,12 for each field. The NDVI maps and NDVI values of each sampling point were represented in Figure 4.10. In contrast to the NDWI values, the maximum of NDVI values was 0.842, the L09 in the Cochno Low, and the minimum was 0.201, the H06 in the Cochno Low.

The NDVI histograms are presented in Figure 4.11. There was no apparent pattern in the distribution of NDVI among fields. Both the histograms and the kernel density curves illustrated that the NDVI values of all samples did not follow a normal distribution, and the Shapiro-Wilk test results proved that ($p<0.001$). The Box-cox transformation was carried out with the NDVI data; but the transformed NDVI were still not normally distributed. Figure 4.12 was the kernel density curve of NDVI, and from the figure, there was the apparent right skewness for NDVI values, except Dumgoyne North. The red line is the mean value of all samples, and the blue line is the median of all samples.

As the NDVI did not follow the normal distribution, the Kruskal-Wallis test was applied to analyse the variation among fields. The null hypothesis was that the NDVI of each field is an identical population. The p-value given by Kruskal-Wallis test turned out nearly zero (3.71e-09). Hence, the null hypothesis was rejected. At the 0.05 significance level, it was concluded that the NDVI values of each field were different populations.

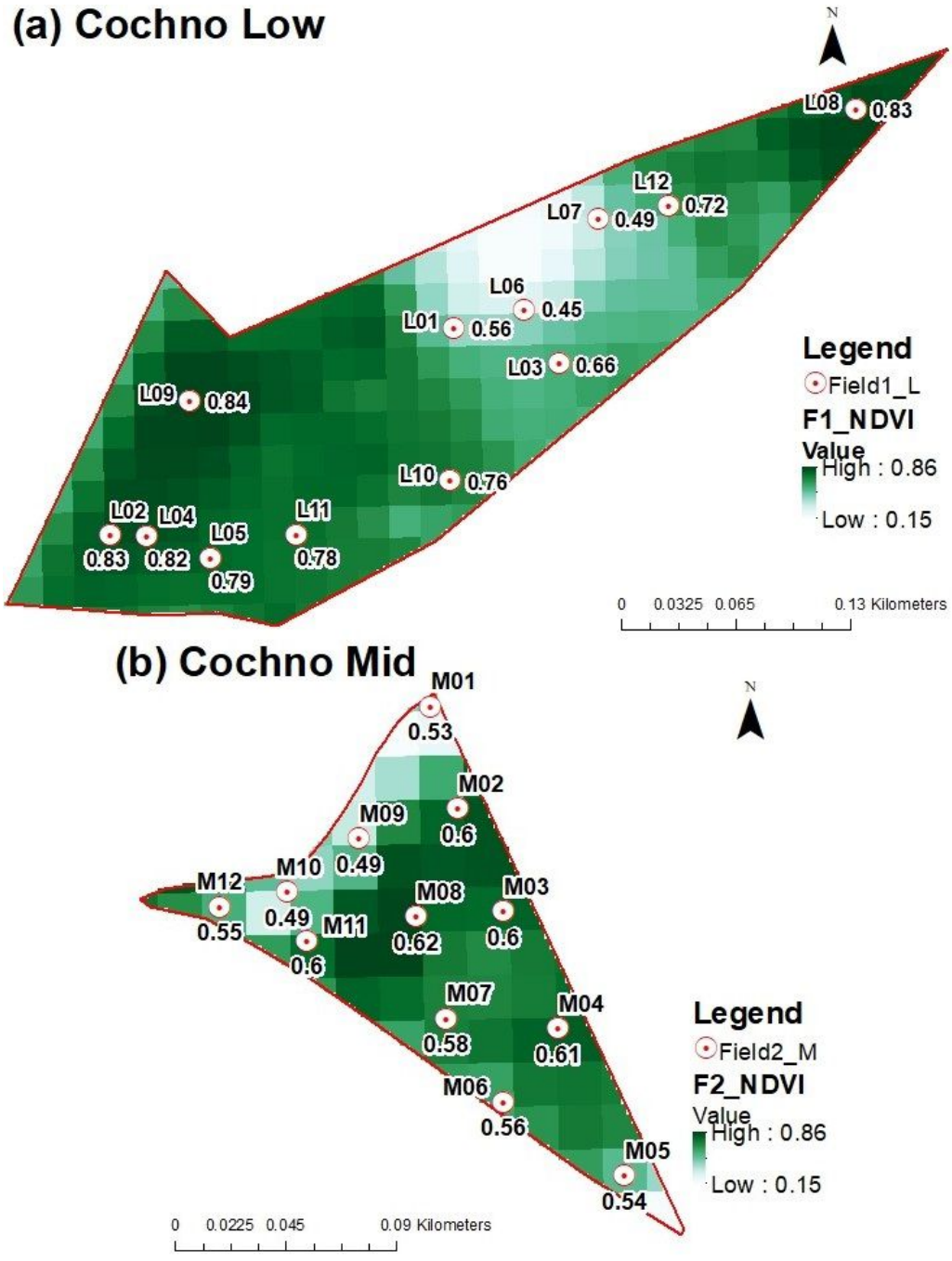
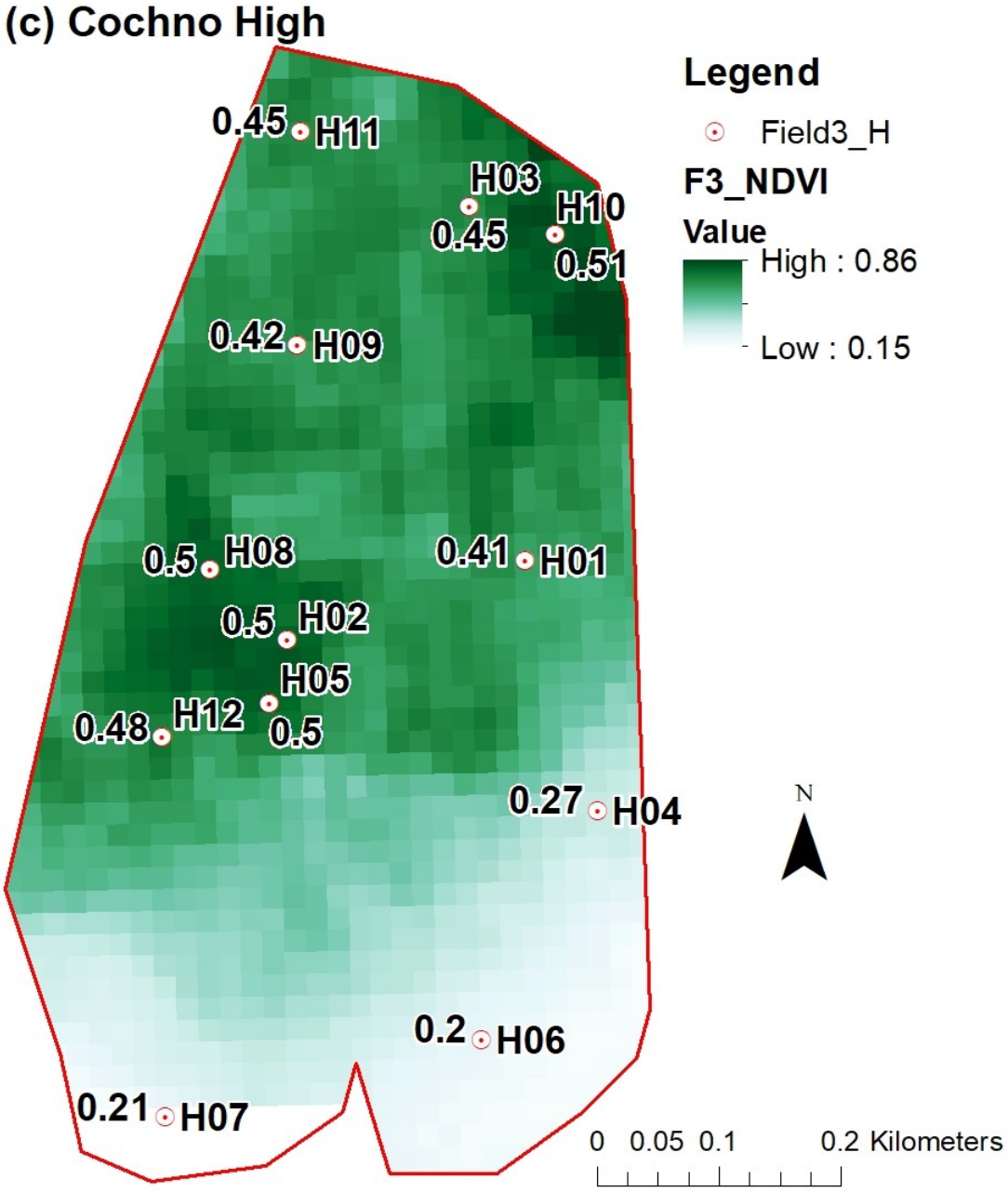
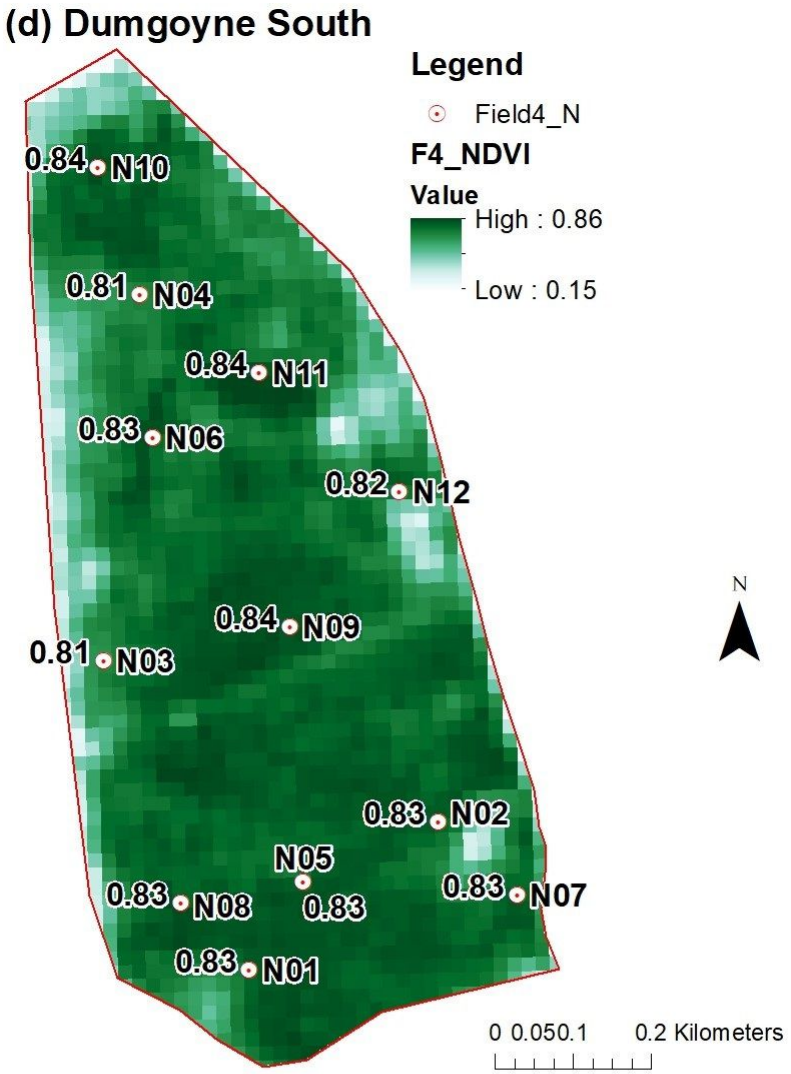
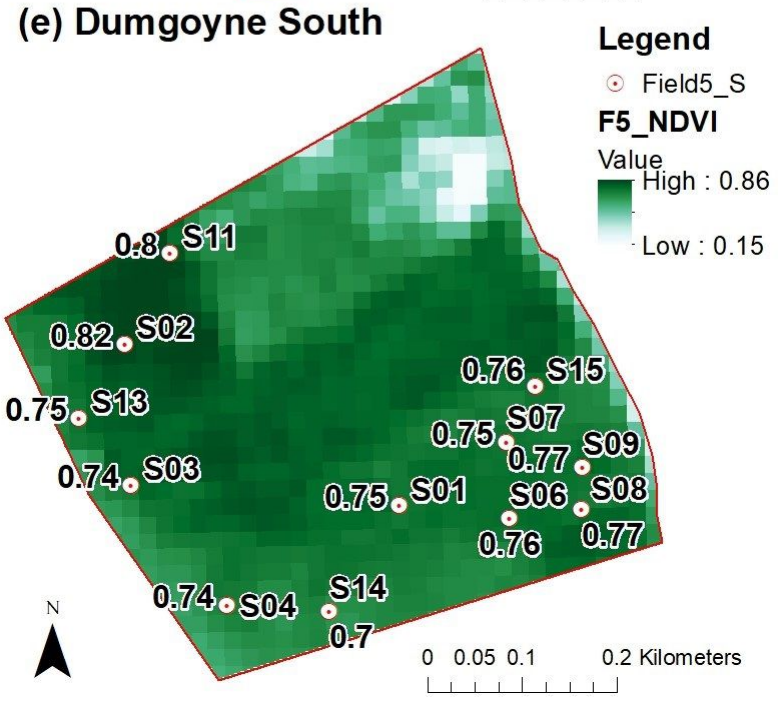


Figure 4.10: The NDVI maps for the study fields and the NDWI of each sampling point: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South.


(c) Cochno High
Legend
Field3_H
F3_NDVI
Value
High : 0.86
Low : 0.15
0.45 H11
H03
0.45
H10
0.51
0.42 H09
0.5 H08
0.41 H01
0.5 H02
H05
0.48 H12
0.5
0.27 H04
N
0.2 H06
0.21 H07
0 0.05 0.1 0.2 Kilometers


(d) Dumgoyne South
Legend
Field4_N
F4_NDVI
Value
High : 0.86
Low : 0.15
0.84 N10
0.81 N04
0.84 N11
0.83 N06
0.82 N12
0.84 N09
0.81 N03
0.83 N02
N05
0.83 N08
0.83
0.83 N07
0.83 N01
N
0 0.050.1 0.2 Kilometers


(e) Dumgoyne South
Legend
Field5_S
F5_NDVI
Value
High : 0.86
Low : 0.15
0.8 S11
0.82 S02
0.76 S15
0.75 S13
0.75 S07
0.77 S09
0.74 S03
0.75 S01
S06
S08
0.76
0.77
0.74 S04
S14
0.7
N
0 0.05 0.1 0.2 Kilometers

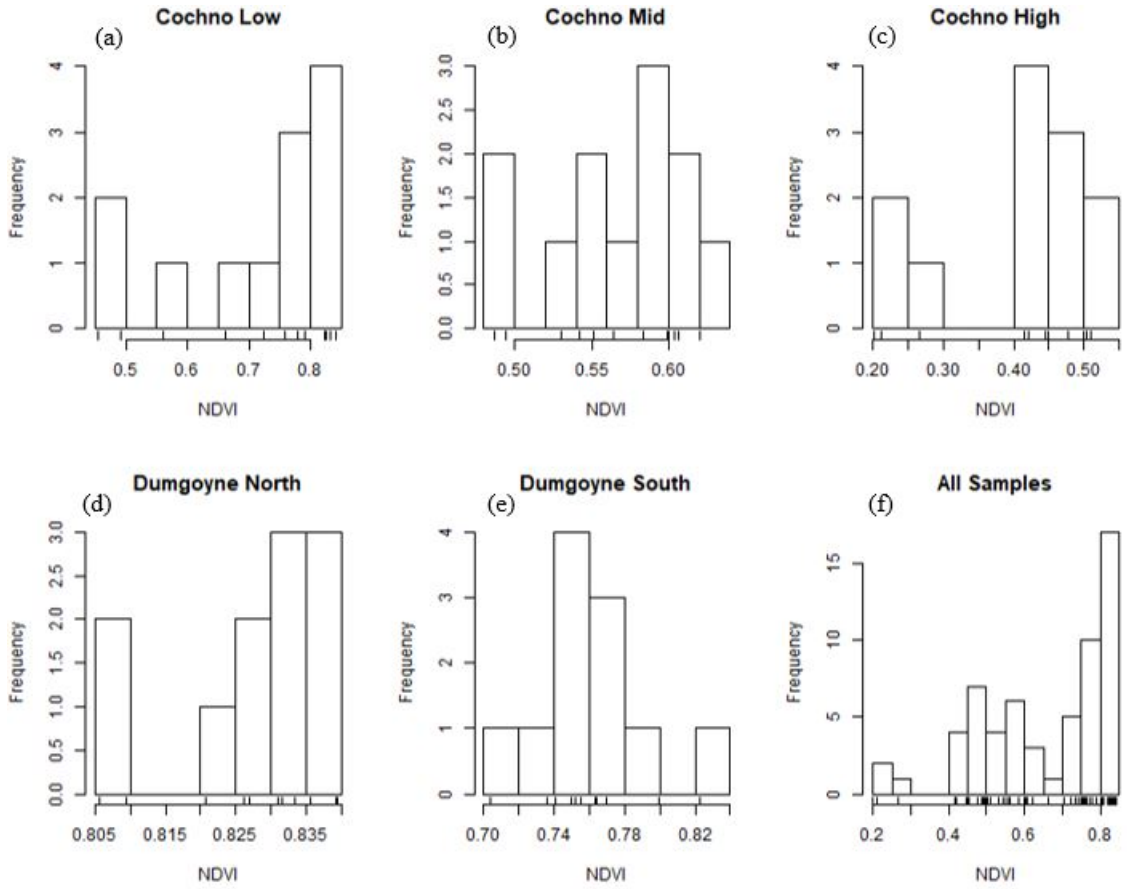


Figure 4.11: Histograms of NDVI for each field and for all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples.

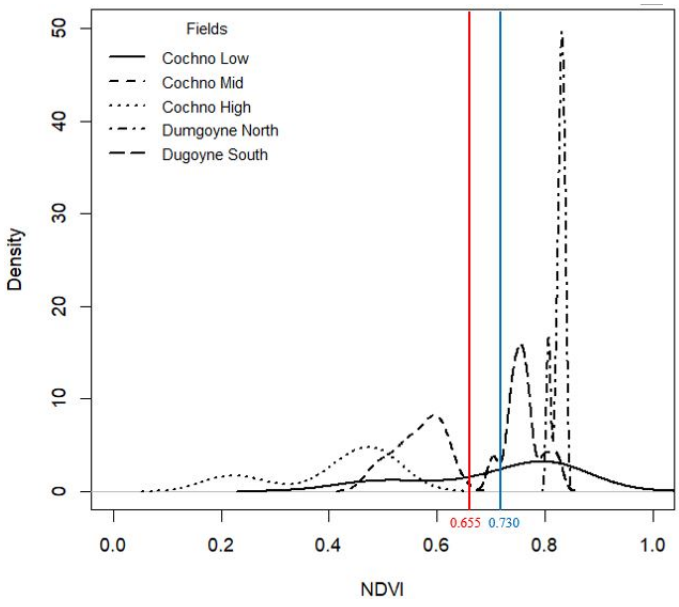


Figure 4.12: Density curve of NDVI for each field.

4.1.3 Topographical factors

4.1.3.1 Elevation

The elevation, which is in WGS84 above sea level, was a consideration when selecting study areas. For this reason, the elevation of sampling locations was significantly different among fields. Cochno has a higher average elevation than Dumgoyne. There was just a slight altitude difference between Dumgoyne fields, whereas the difference in altitude between Cochno Mid and Cochno High is nearly 120 m. Figure 4.13 is the elevation map of each field.

Figure 4.14 displays the histograms for the elevation of each field. The Shapiro-Wilk normality test showed that there was a non-normal distribution of the elevation ($p<0.001$). The kernel density curves in Figure 4.15 illustrated the variation of elevation among fields. There were clear boundaries in-between the density curves of the Cochno fields, which corresponds to the altitude differences among fields. However, the elevation of each field showed a very high density within a limited altitude range, which means there was no significant difference among sampling locations within each field. From the kernel density curve and histogram, the elevation data showed the variation among fields. The Kruskal-Wallis test was carried out to examine whether this variation was statistically or not. The result was given by the R function, and the p-value was 5.255E-11. Hence, we can conclude the difference of elevation by fields was significant.

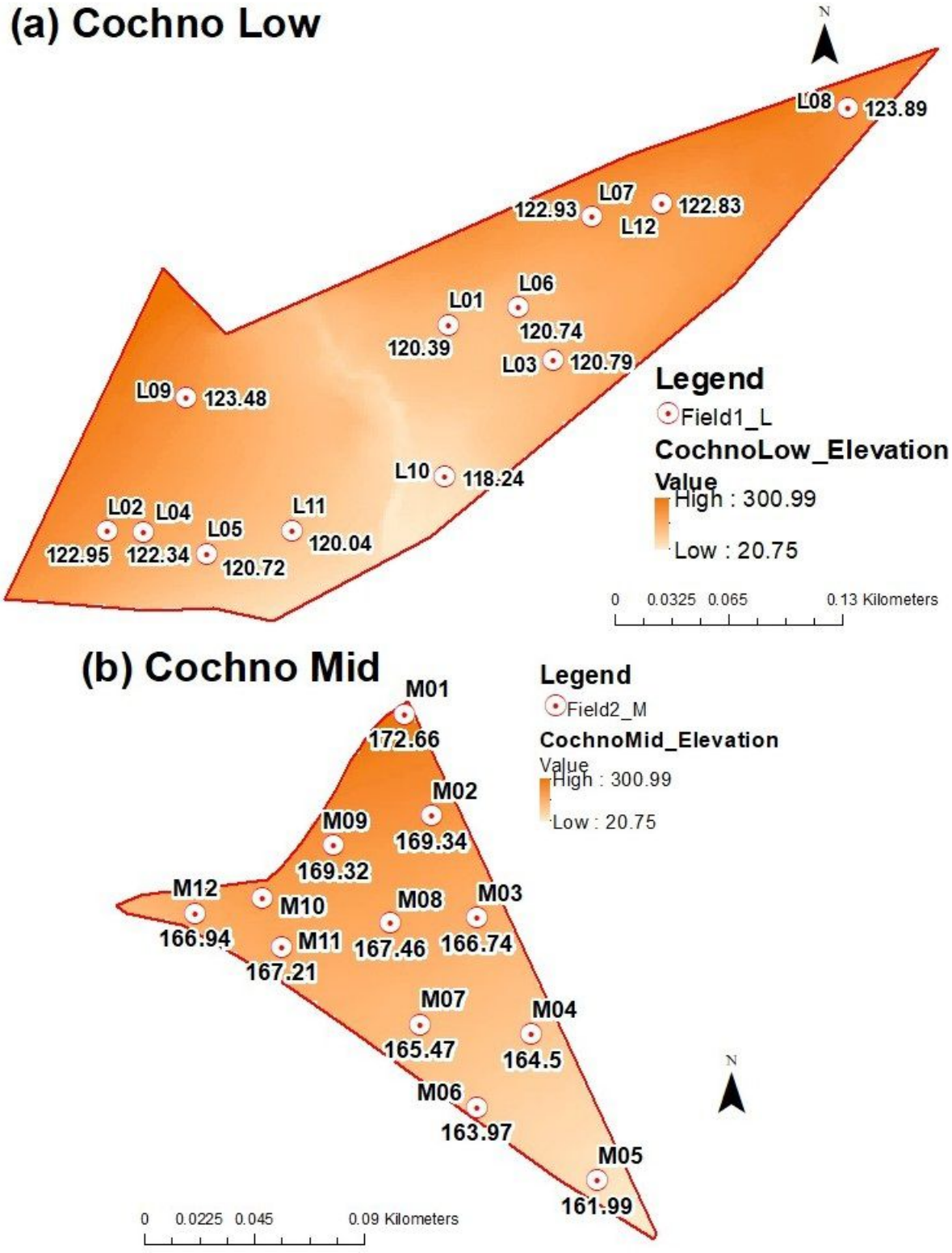
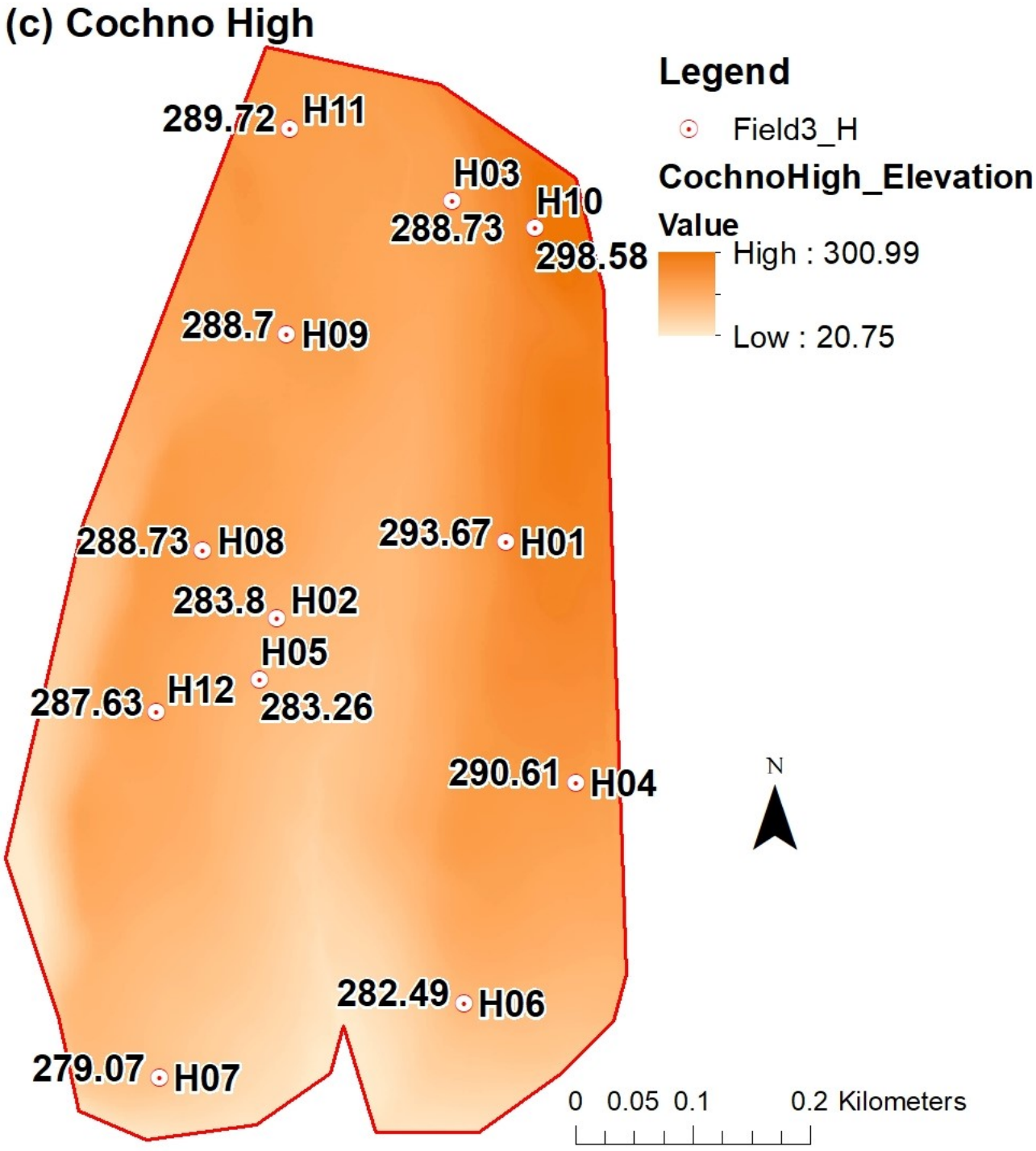
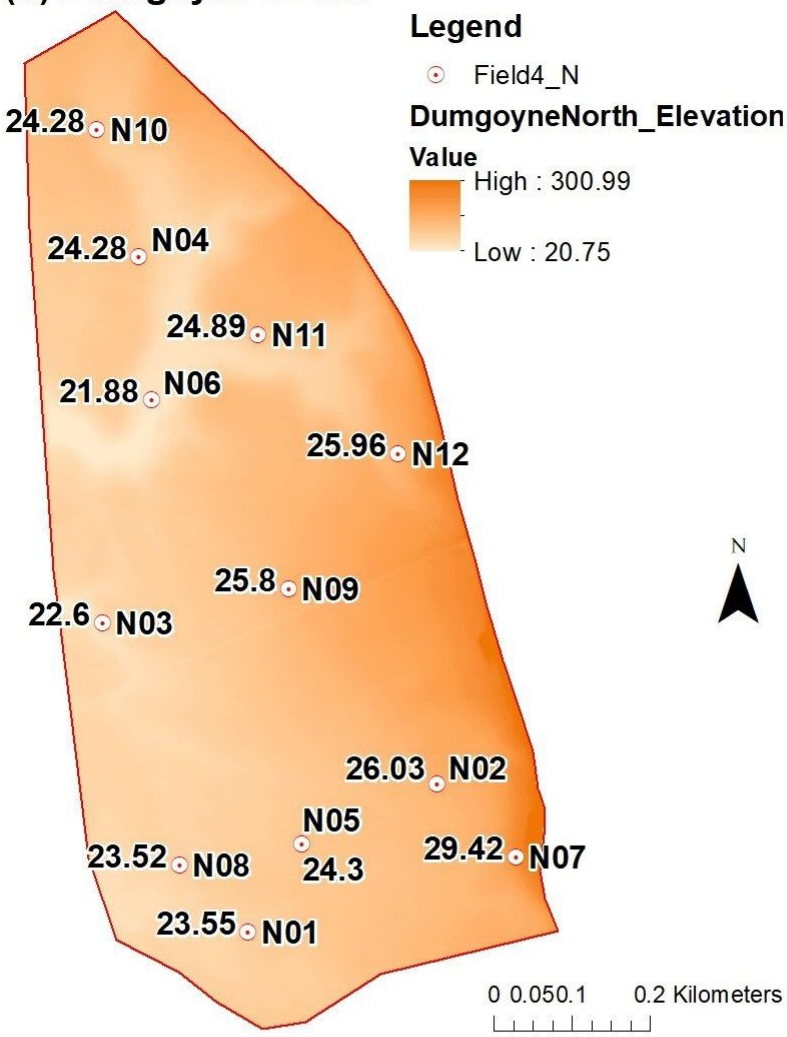


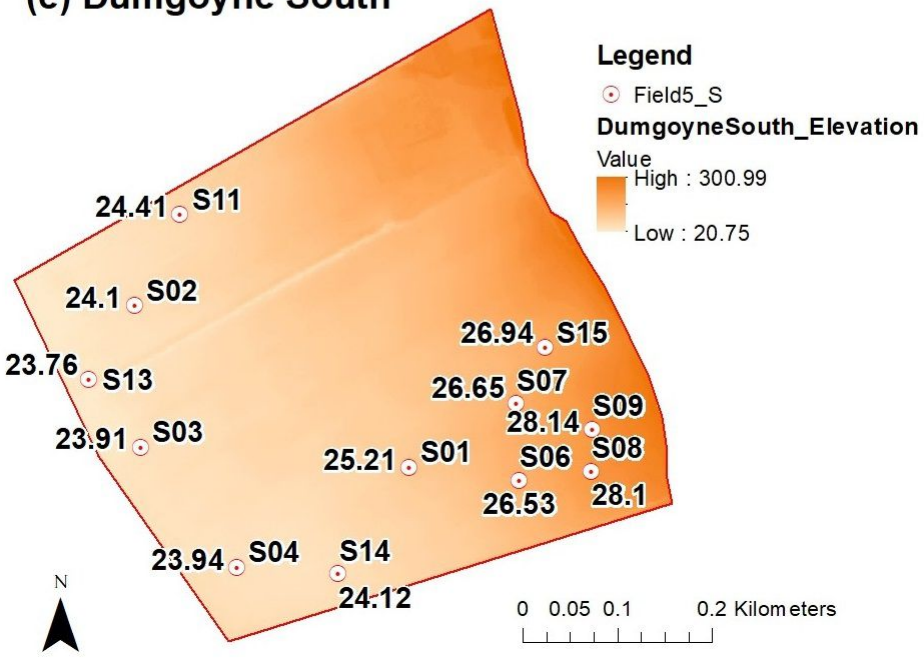
Figure 4.13: The elevation maps for the study fields and the altitude of each sampling point: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South.


(c) Cochno High
Legend
Field3_H
CochnoHigh_Elevation
Value
High : 300.99
Low : 20.75
289.72 H11
H03
H10
288.73
298.58
288.7 H09
288.73 H08
293.67 H01
283.8 H02
H05
287.63 H12
283.26
290.61 H04
282.49 H06
279.07 H07
N
0 0.05 0.1 0.2 Kilometers

(d) Dumgoyne South



(e) Dumgoyne South



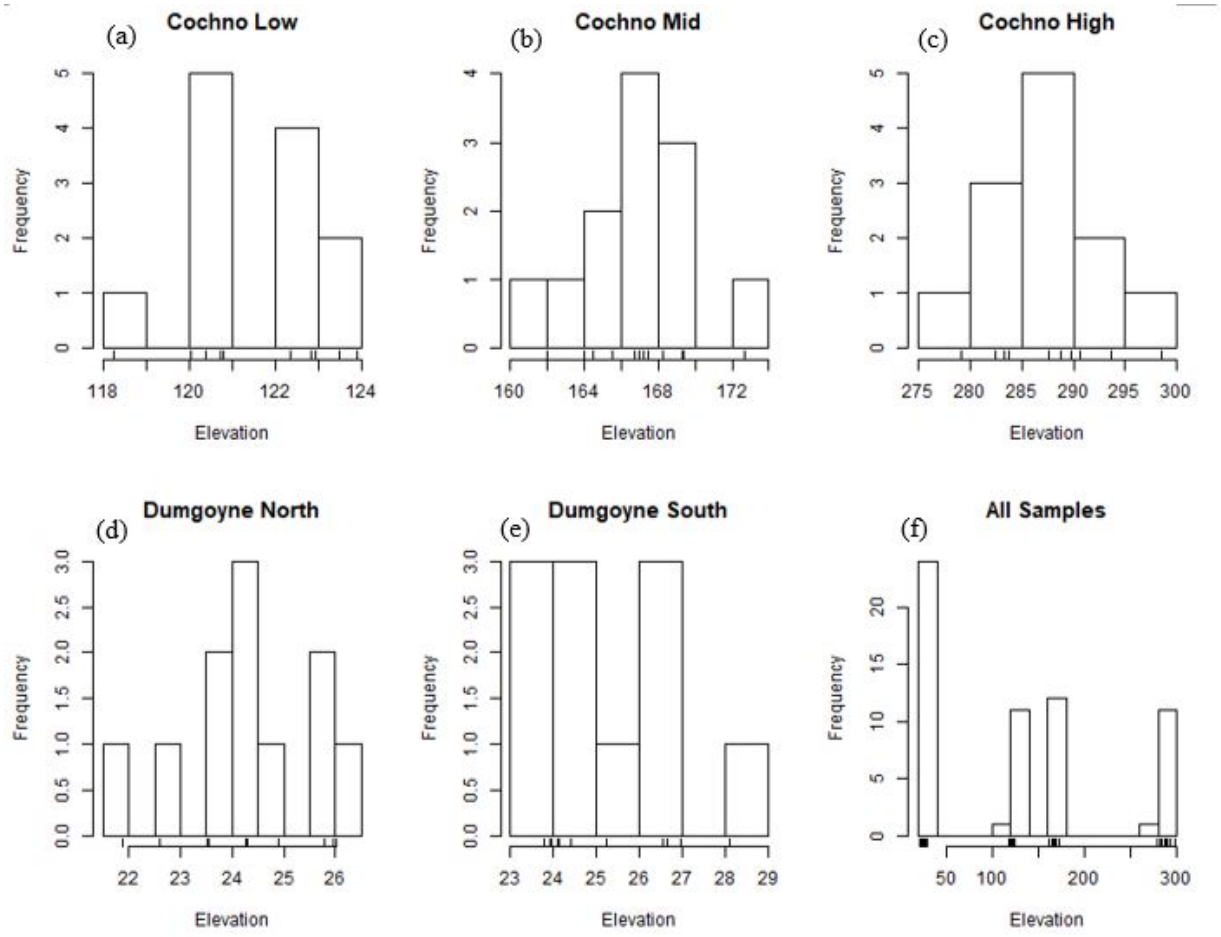


Figure 4.14: Histograms of elevation for each field and all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples.

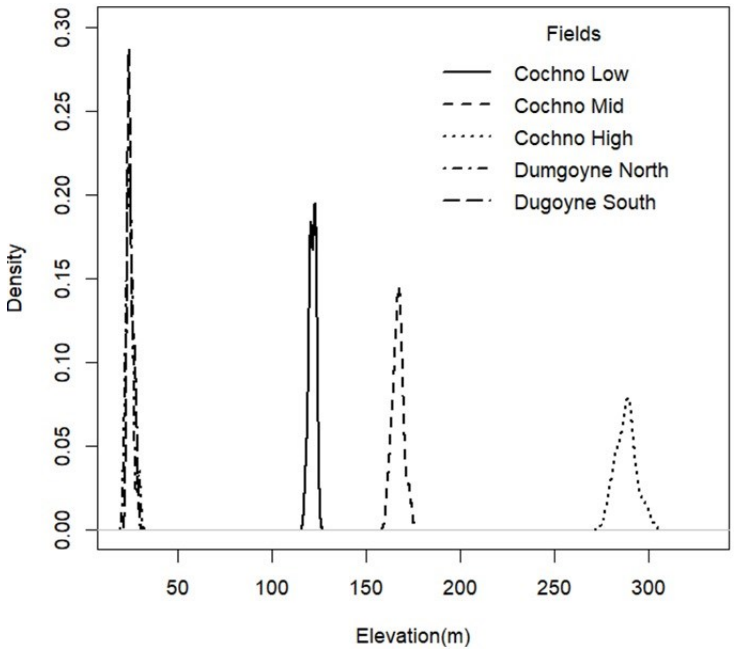


Figure 4.15: Density curve of elevation for each field.

4.1.3.2 Slope

The slope values of sampling points were obtained from the slope map, calculated from the DEM images. Figure 4.16 displays the slope maps, and the slope value of each sampling point is also labelled. From the maps, we can see that Cochno Low was flatter than other fields and the terrain of Cochno High was quite undulating. The H01 in Cochno High had the steepest slope (10.4 degrees), and 0.6 degrees was the lowest slope (S04 in Dumgoyne South).

The histograms (Figure 4.17) illustrated that the slope was not normally distributed. The red line in Figure 4.18 4.18 is the mean of the slope, and the blue is the median of the slope. The mean of samples was larger than the median, which suggested that the data was highly skewness to the right. The result of the Shapiro-Wilk normality test demonstrated the data did not follow the normal distribution. After the Box-Cox transformation, the transformed Slope converted to the normal distribution. The Homogeneity of variance was checked by Bartlett's test. The variance of each field was statistically equal (p-value is 0.8215 in Bartlett's test). The One-way analysis of variance was applied to check the variation of the slope by field, and the p-value of F-test was 0.000000119. At the 95% Confidence interval level, we rejected the null hypothesis and admitted that at least one field was significantly different with other fields. Moreover, the further analysis showed that the slope of Dumgoyne South was different from other fields.

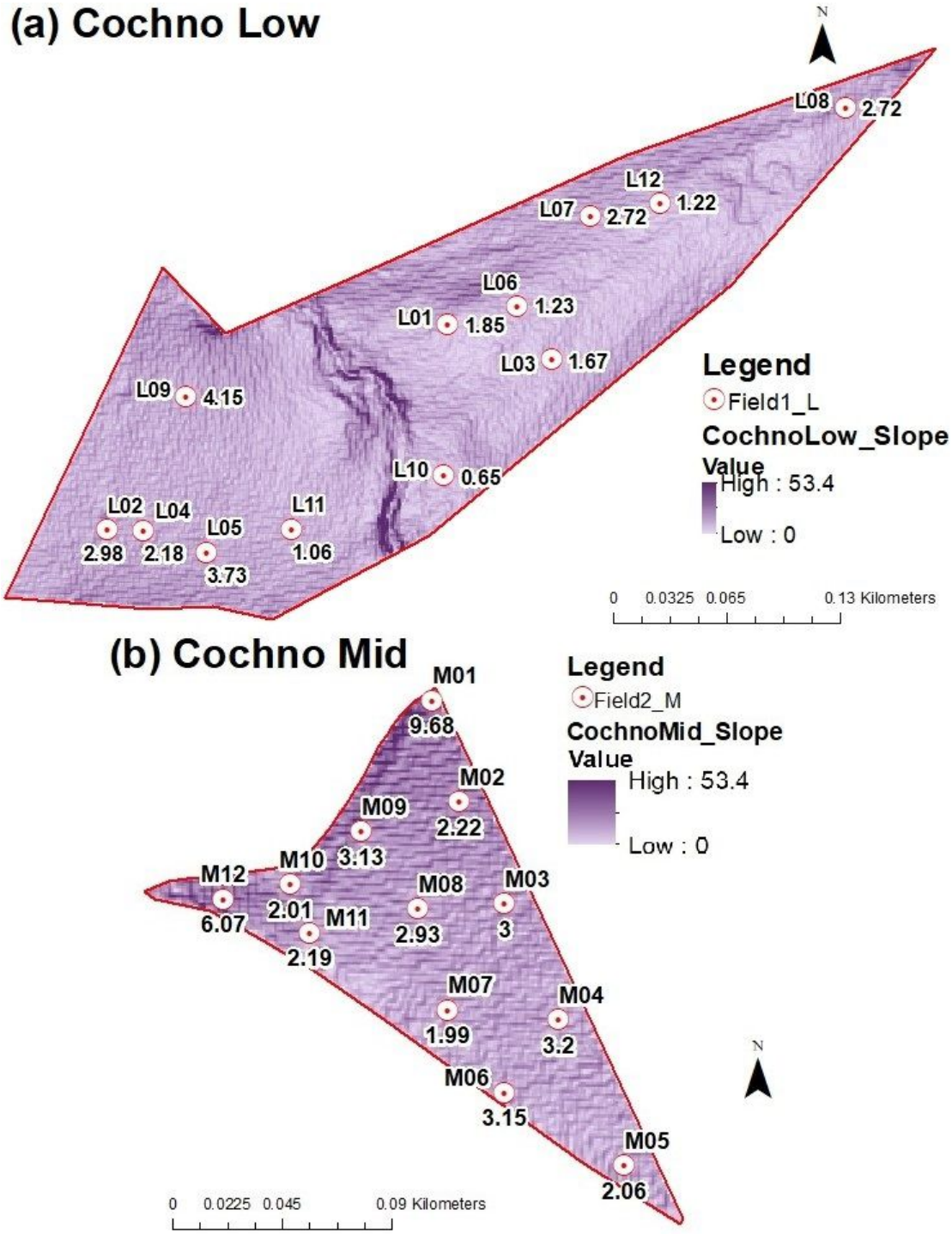
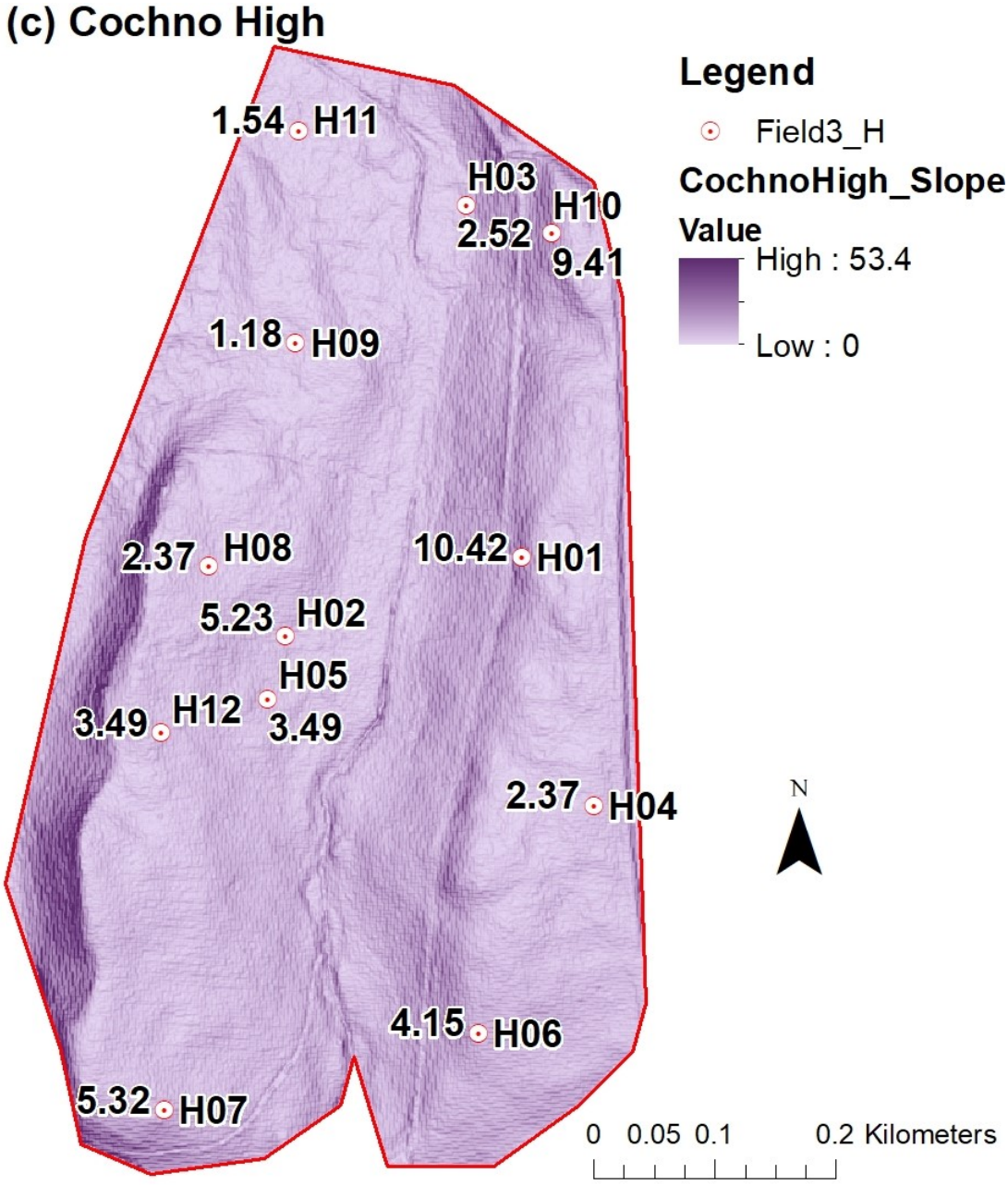
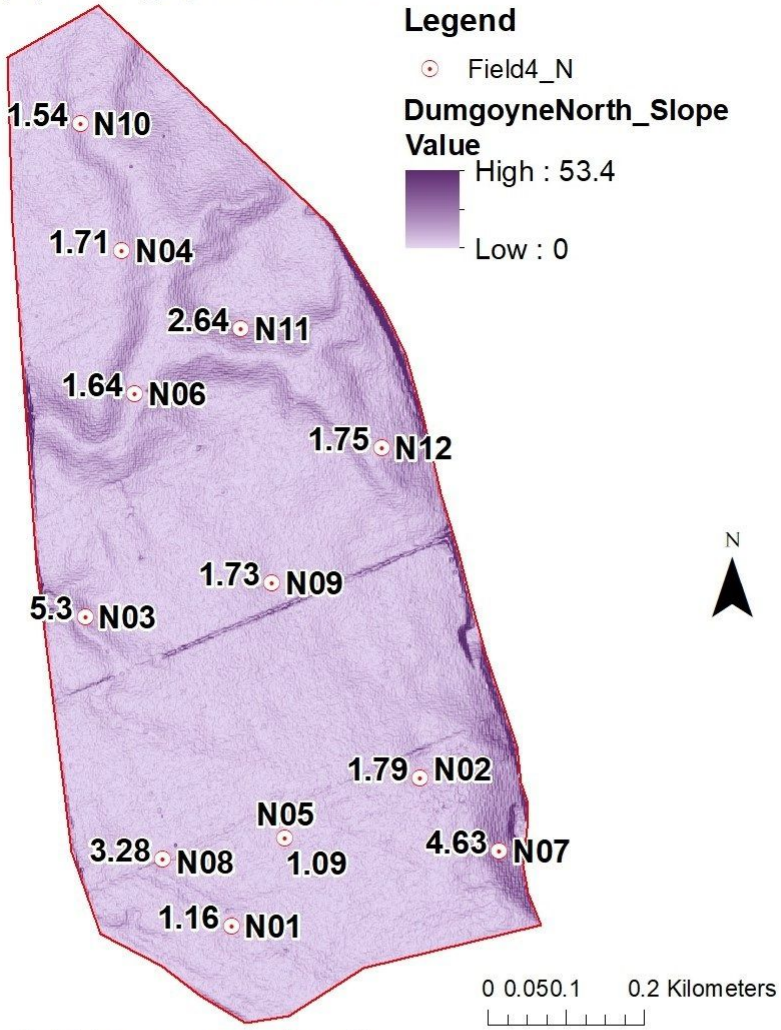


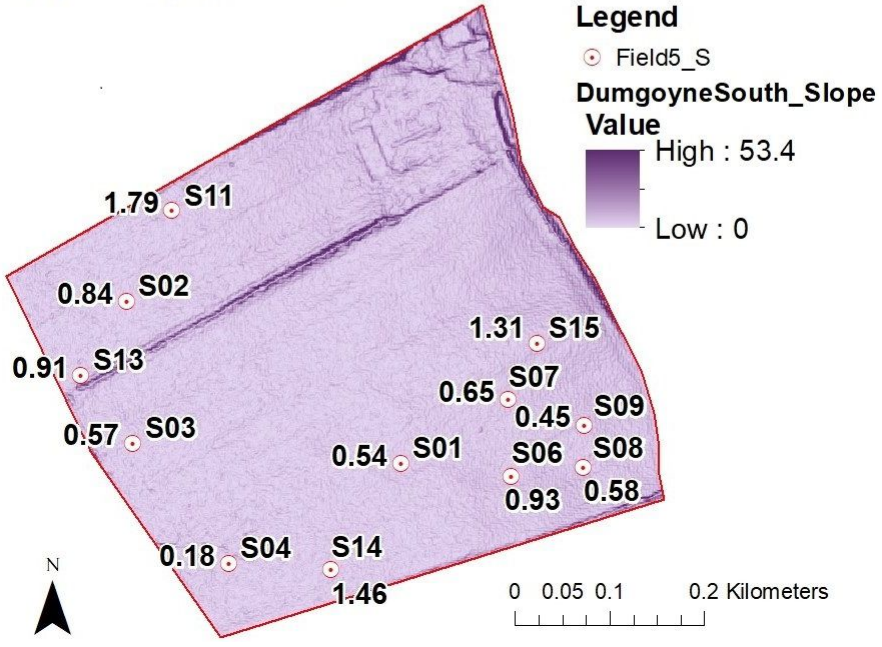
Figure 4.16: The slope maps for the study fields and the slope of each sampling point: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South.


(c) Cochno High
Legend
Field3_H
CochnoHigh_Slope
Value
High : 53.4
Low : 0
1.54 H11
H03
H10
2.52
9.41
1.18 H09
2.37 H08
10.42 H01
5.23 H02
H05
3.49 H12
3.49
2.37 H04
4.15 H06
5.32 H07
N
0 0.05 0.1 0.2 Kilometers

(d) Dumgoyne South



(e) Dumgoyne South



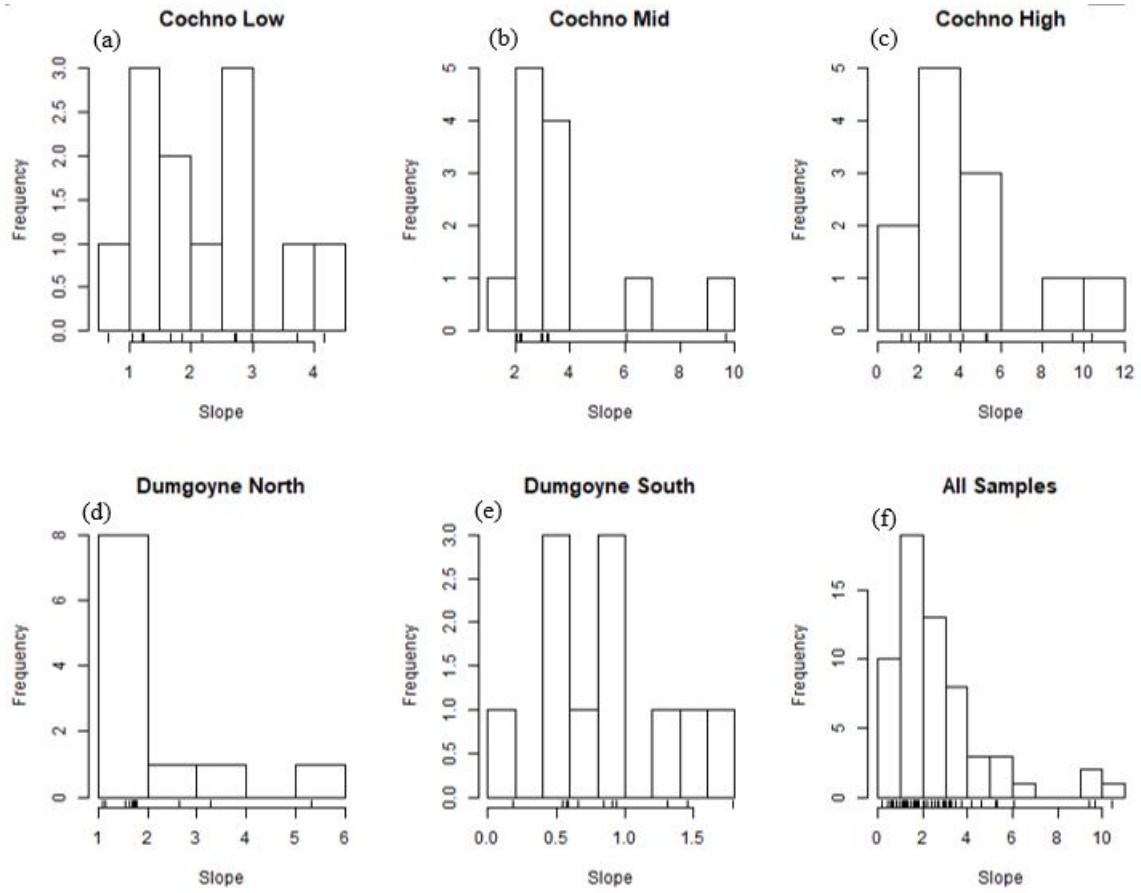


Figure 4.17: Histograms of slope for each field and all sample: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples.

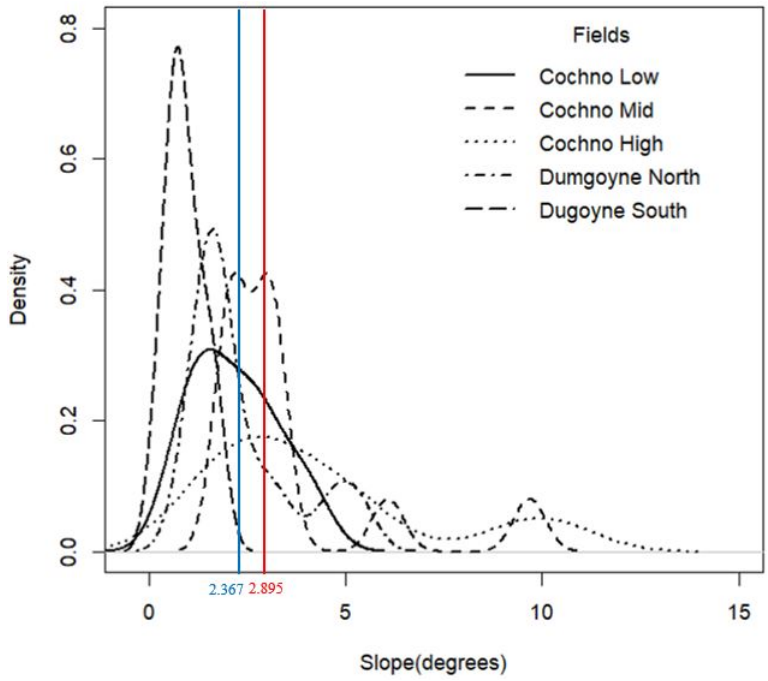


Figure 4.18: Density curve of slope for each field.

4.1.3.3 Aspect

Figure 4.20 displays the aspect maps for each sampling point and Figure 4.19 displays the histogram for aspect in each field. Southeast and Southwest orientations occur most frequently in the study fields, and none of the sampling points face north, south or west.

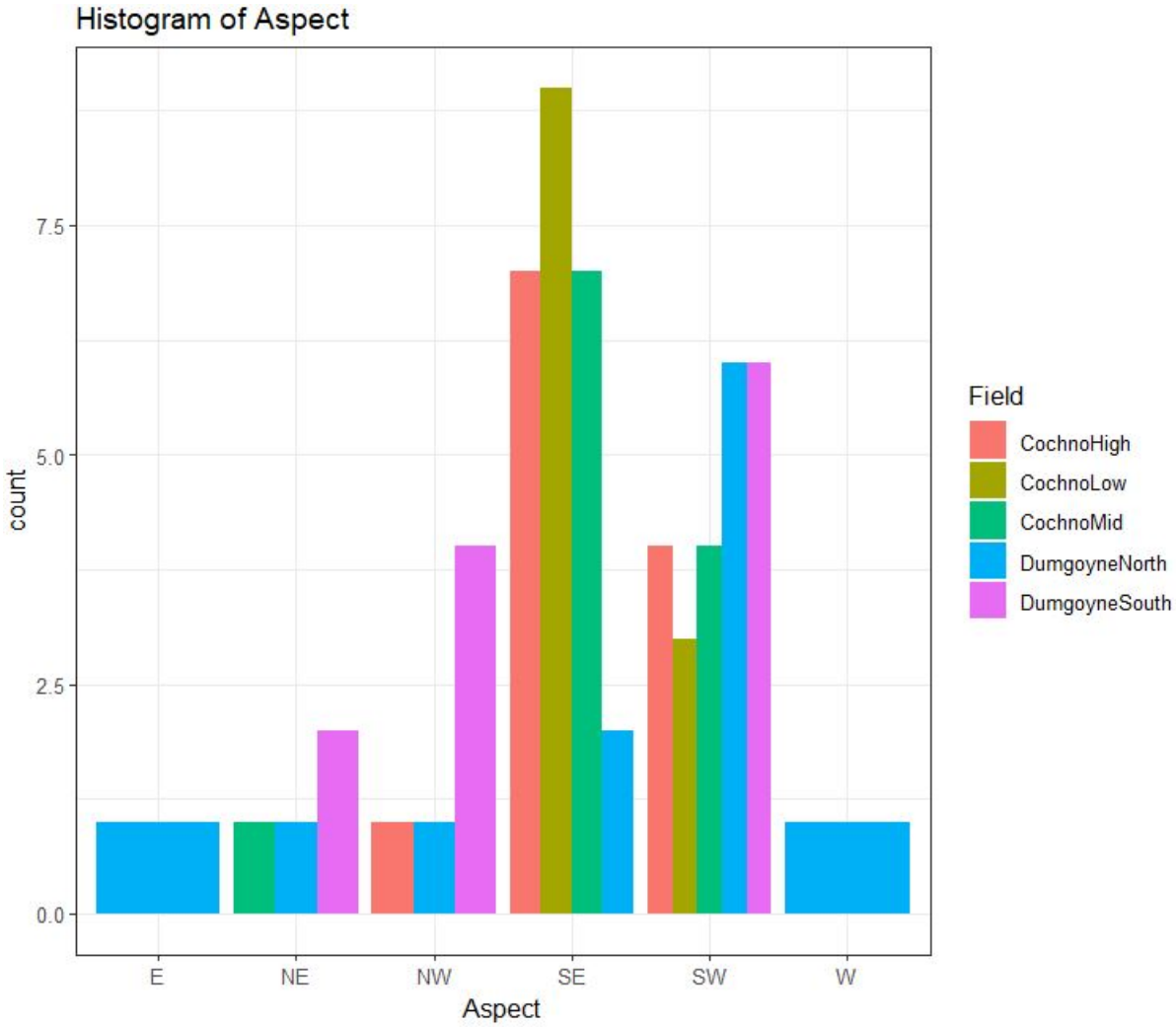


Figure 4.19: Histogram of aspect for each field.

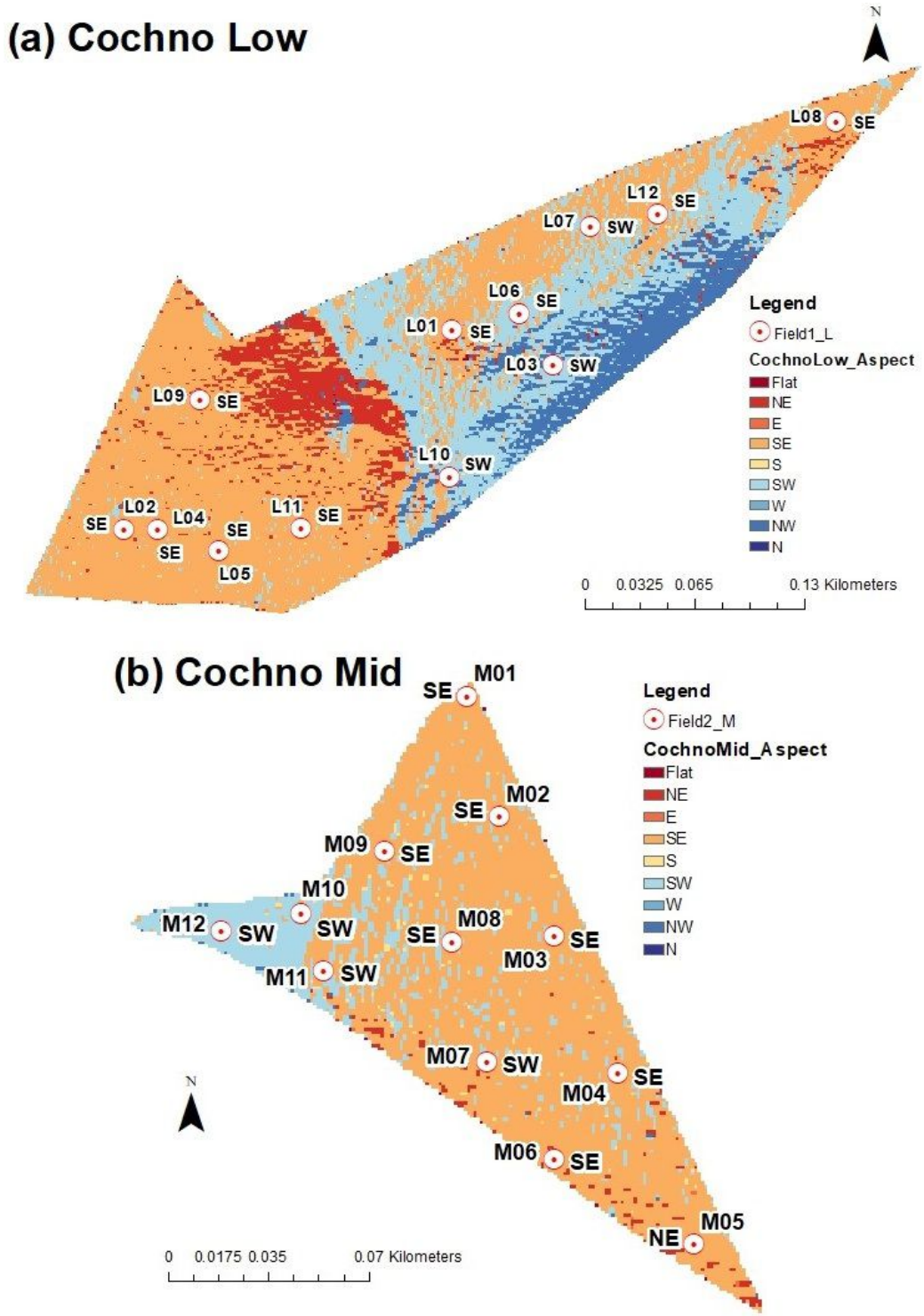
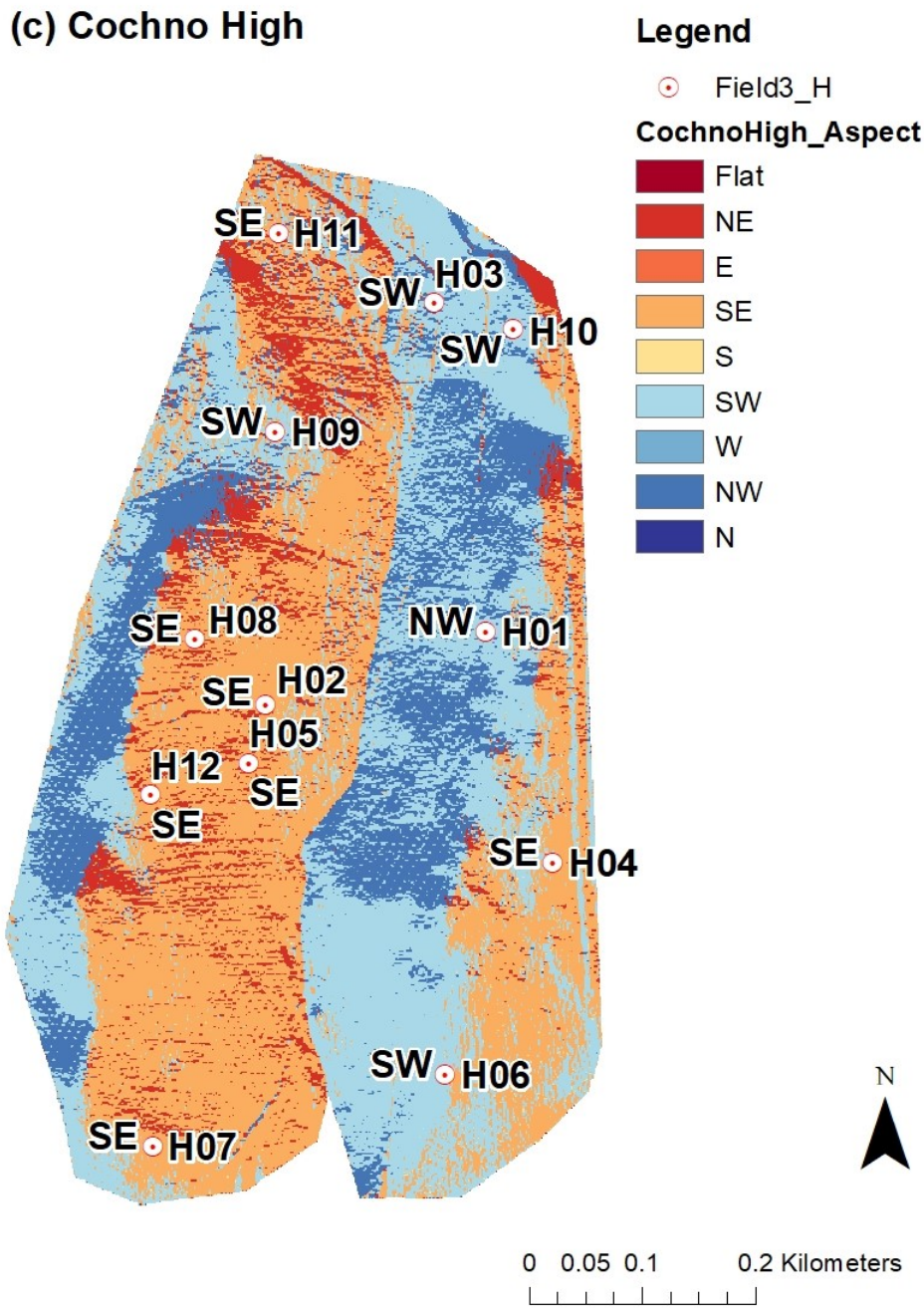
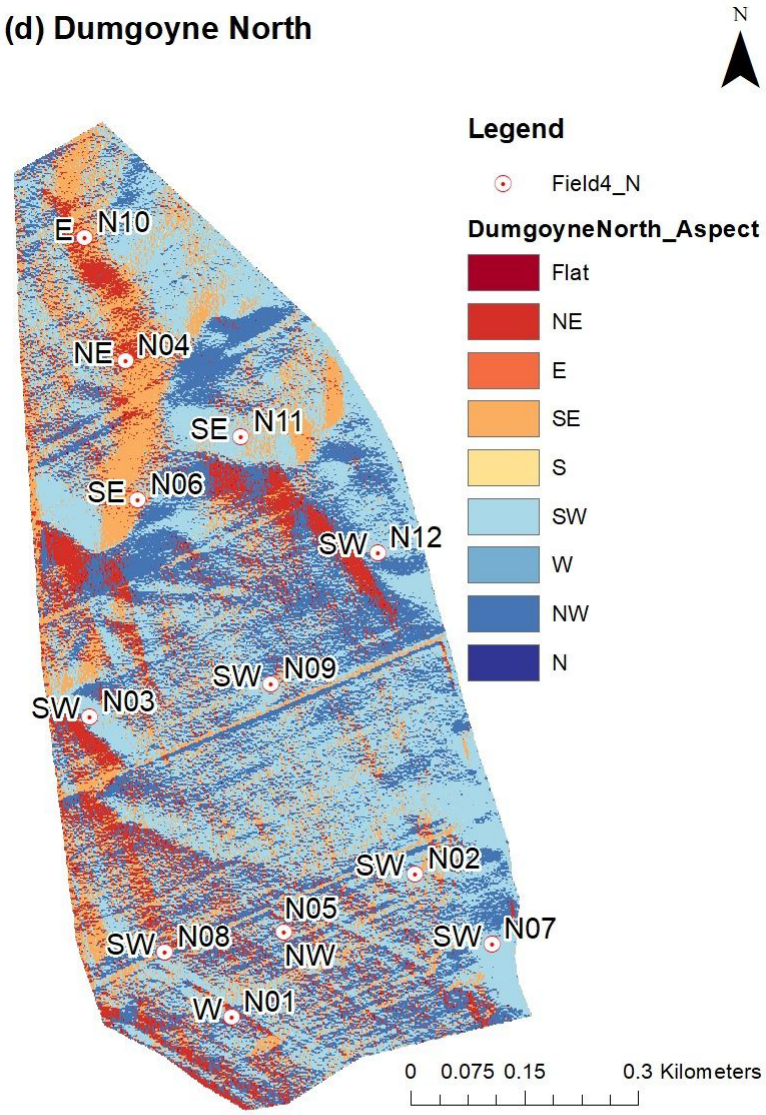
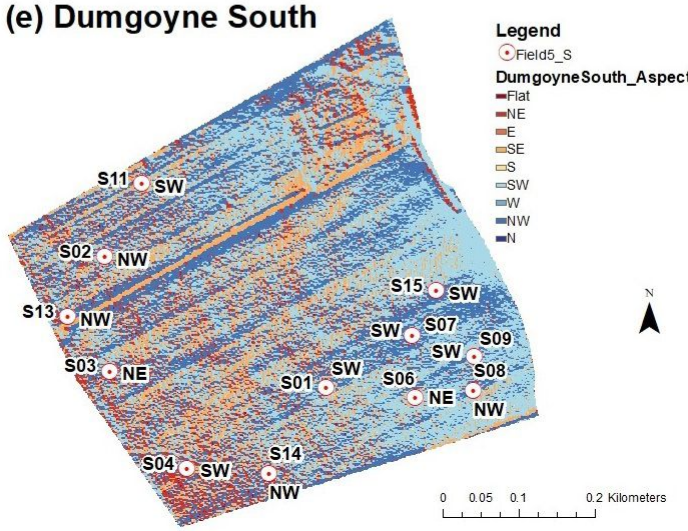


Figure 4.20: The aspect maps for the study fields and the aspect of each sampling point: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South.


(c) Cochno High
Legend
Field3_H
CochnoHigh_Aspect
Flat
NE
E
SE
S
SW
W
NW
N
SE H11
SW H03
SW H10
SW H09
SE H08
NW H01
SE H02
H05
H12
SE
SE
SE H04
SW H06
SE H07
N
0 0.05 0.1 0.2 Kilometers


(d) Dumgoyne North
N
Legend
Field4_N
DumgoyneNorth_Aspect
Flat
NE
E
SE
S
SW
W
NW
N
E N10
NE N04
SE N11
SE N06
SW N12
SW N09
SW N03
SW N02
N05
SW N08
NW
SW N07
W N01
0 0.075 0.15 0.3 Kilometers


(e) Dumgoyne South
Legend
Field5_S
DumgoyneSouth_Aspect
Flat
NE
E
SE
S
SW
W
NW
N
S11 SW
S02 NW
S13 NW
S03 NE
S15 SW
SW S07
S09
SW
S08
S01 SW
S06 NE
NW
S04 SW
S14
NW
0 0.05 0.1 0.2 Kilometers

4.1.4 Soil properties

4.1.4.1 Soil Moisture

The geographical distribution of soil moisture values in each field is represented in (Figure 4.21). Fifty-nine soil moisture values were acquired from the field survey. The point N07 in Dumgoyne North was not measured. The highest soil moisture value appeared in Cochno High, where the soil moisture of the H01 was 0.744, and the lowest value was 0.435 (M12). The average soil moisture of all sample points was 0.588. The histogram plots illustrated that the soil moisture values of each field were differently distributed within a certain range. These distributions did not follow a clear pattern, and there were differences among these distributions. For example, most fields had the highest frequency at the high values (above 0.5 or 0.6), despite Cochno High, in which the highest frequency was at 0.4 to 0.5. This was also obviously shown in the kernel density curve (Figure 4.23). The dotted line, the density curve of Cochno High, was right-skewed while other curves were left-skewed. The mean of all samples is nearly equal to the median, which suggested the data was less skewness. Although the soil moisture data did not follow the normal distribution according to the result of the Shapiro-Wilk normality test, they were able to transform to a normal distribution by the Box-Cox transformation. Also, the modified values did not show the heterogeneity of different fields. After the previous analysis, a one-way ANOVA was conducted to find out the difference among fields. The p-value of F-test was 0.009 , less than 0.05. Therefore, the mean of soil moisture in each field was not equal. There was at least one field having the significant difference with other fields.

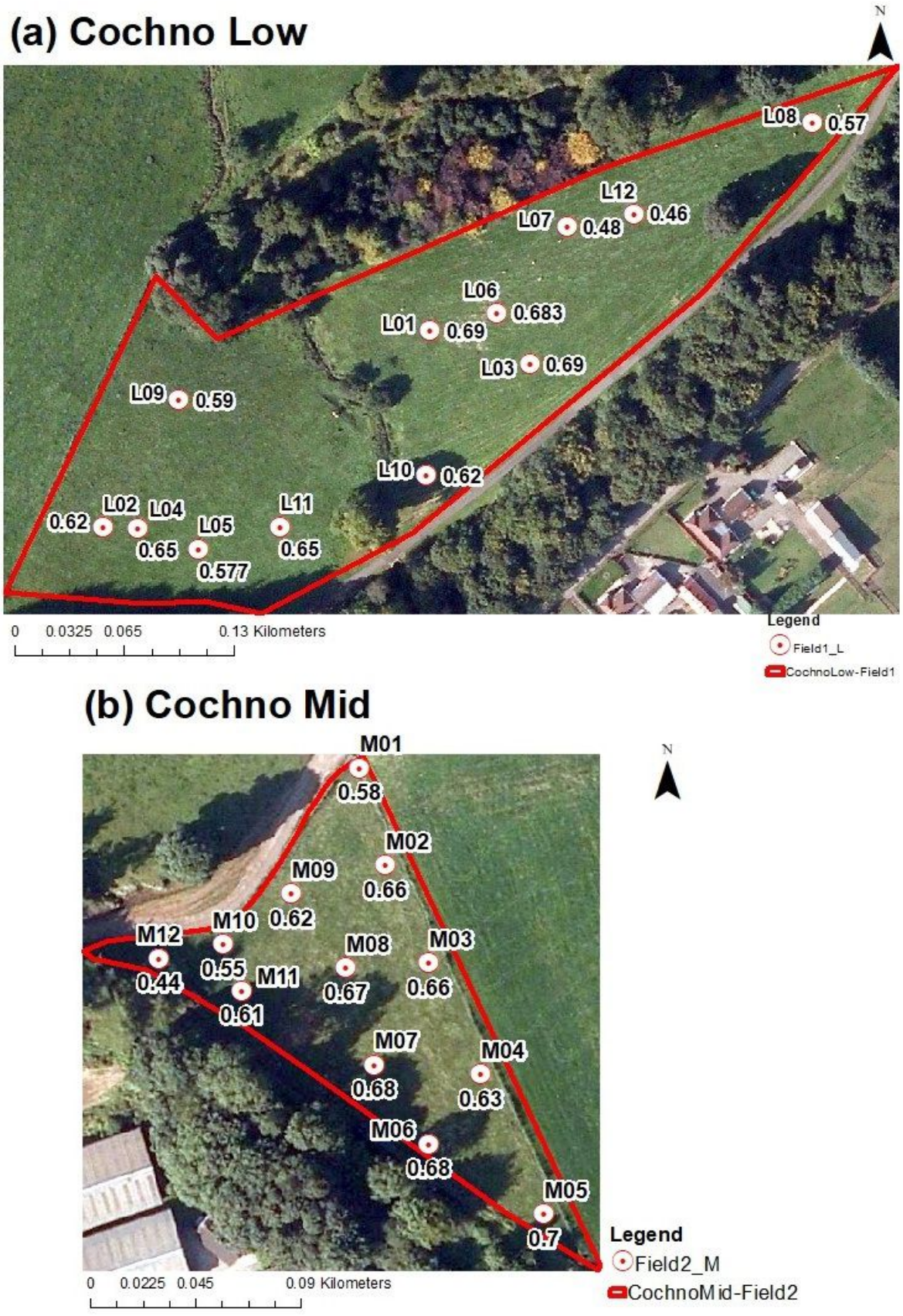
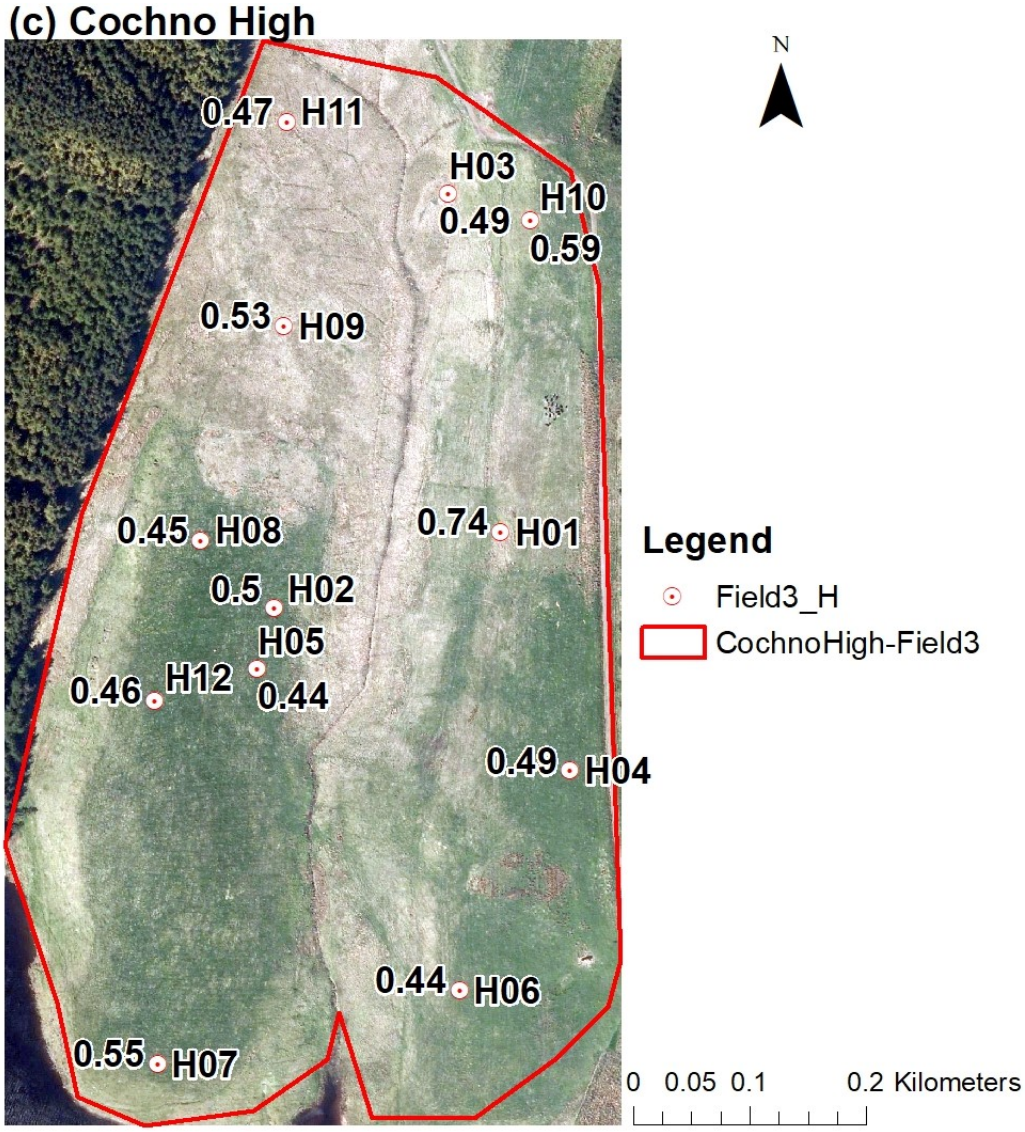
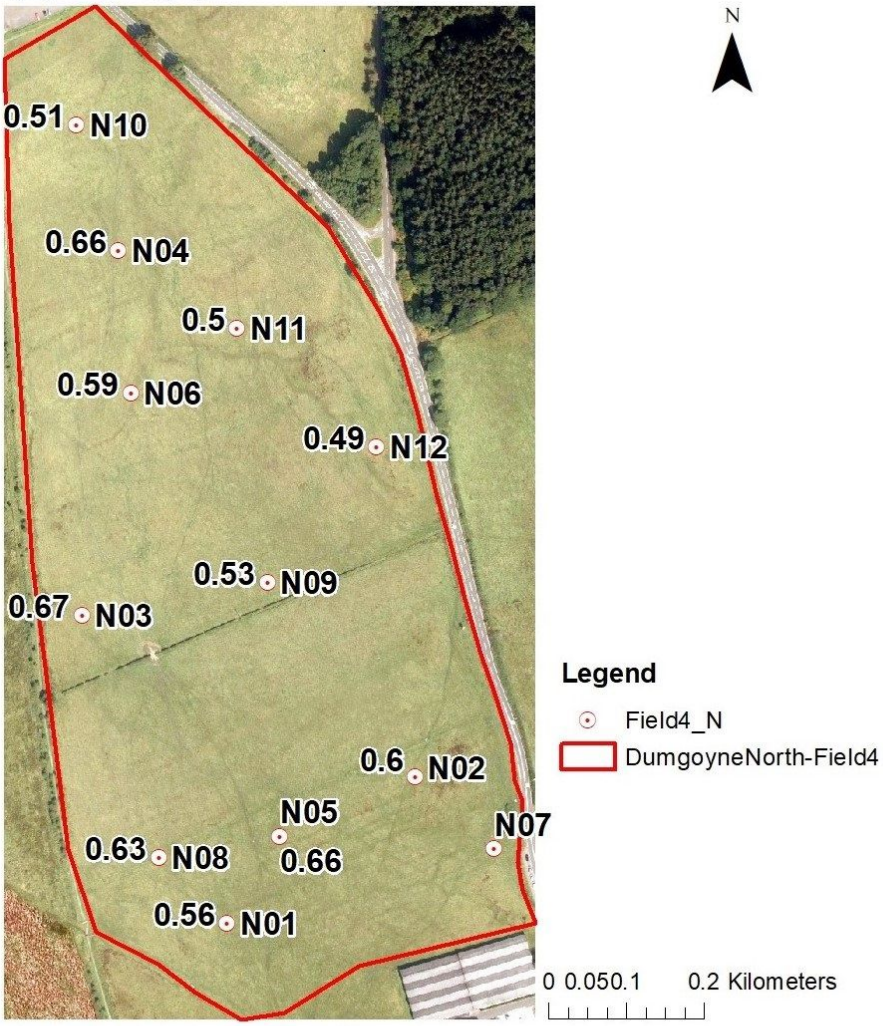


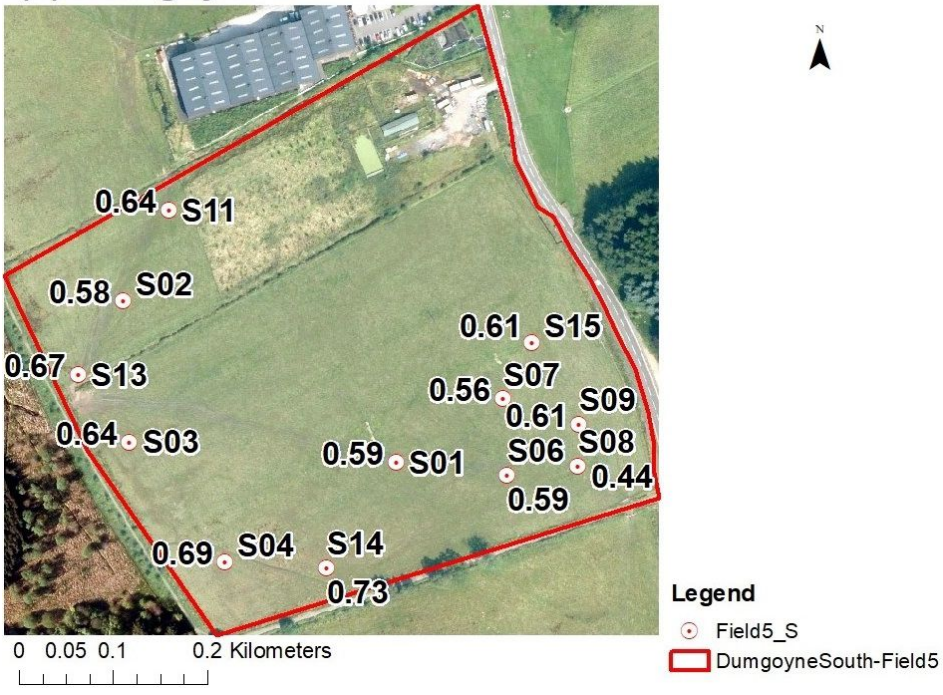
Figure 4.21: The soil moisture values for sampling points within the study fields: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South.


(c) Cochno High
N
0.47 H11
H03
H10
0.49
0.59
0.53 H09
0.45 H08
0.74 H01
0.5 H02
H05
0.46 H12
0.44
0.49 H04
0.44 H06
0.55 H07
Legend
Field3_H
CochnoHigh-Field3
0 0.05 0.1 0.2 Kilometers

(d) Dumgoyne South



(e) Dumgoyne South



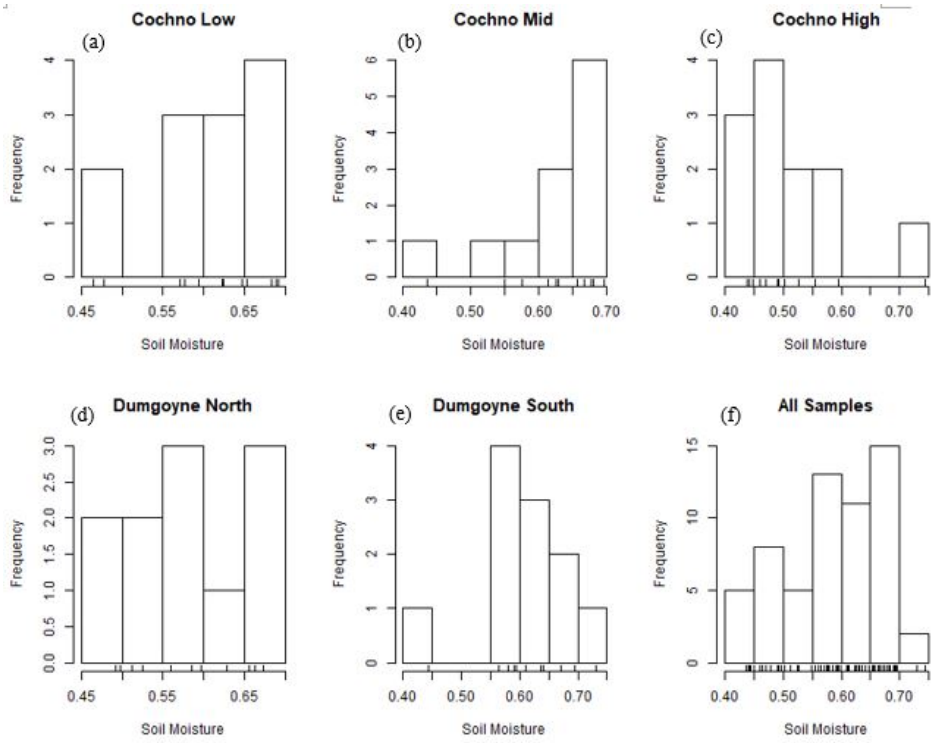


Figure 4.22: Histograms of soil moisture for each field and all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples.

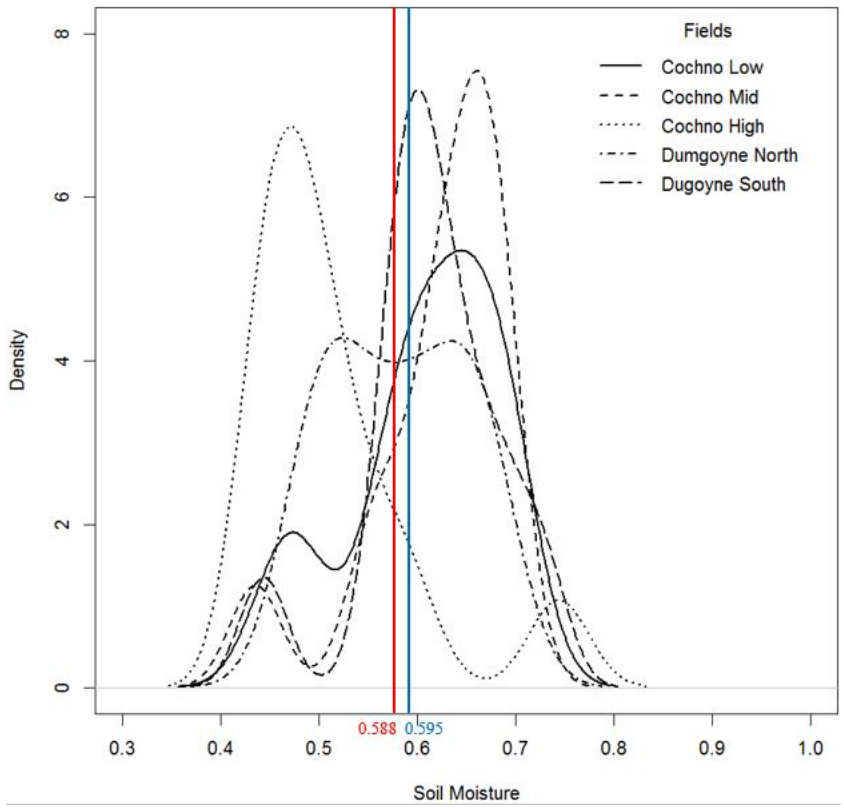


Figure 4.23: Density curve of soil moisture for each field.

4.1.4.2 Soil Temperature

Figure 4.24 displays the soil temperature values in each field. Similar to the soil moisture collection, 59 samples were recorded. The highest recorded soil temperature was 19 degrees in the N09 (Dumgoyne North), and the lowest soil temperature was 12 degrees in the H01. H01 had the highest soil temperature and soil moisture among all sampling points.

Figure 4.25 contains the histogram for each field and also for all samples. There was a high frequency of soil temperature values in the range of 12°C to 16°C. It was evident that the distribution of soil temperature values varied among fields. The kernel density curve of each field in Figure 4.26 also supported this. The red line in Figure 4.26 represents the average soil temperature value for all samples and the blue line indicates the median values of all samples. From the density curve, we can see that Dumgoyne South (long dash line) and Cochno Mid (short dash line) were similar in distribution with a left-skewed curve. By contrast, Cochno Mid density curve (dotted line) and Cochno Low curve (solid line) showed different distributions. Furthermore, the soil temperature values were not normally distributed and unable to transform into the normal distribution. The Kruskal-Wallis test was carried out to detect the variation of soil temperature by fields. The result is displayed in Appendix D. The p-value (<0.001) of the Kruskal-Wallis test was nearly zero which suggested that the soil temperature values by fields did not have an equal mean, and therefore we can conclude that the variation of soil temperature among fields was significant.

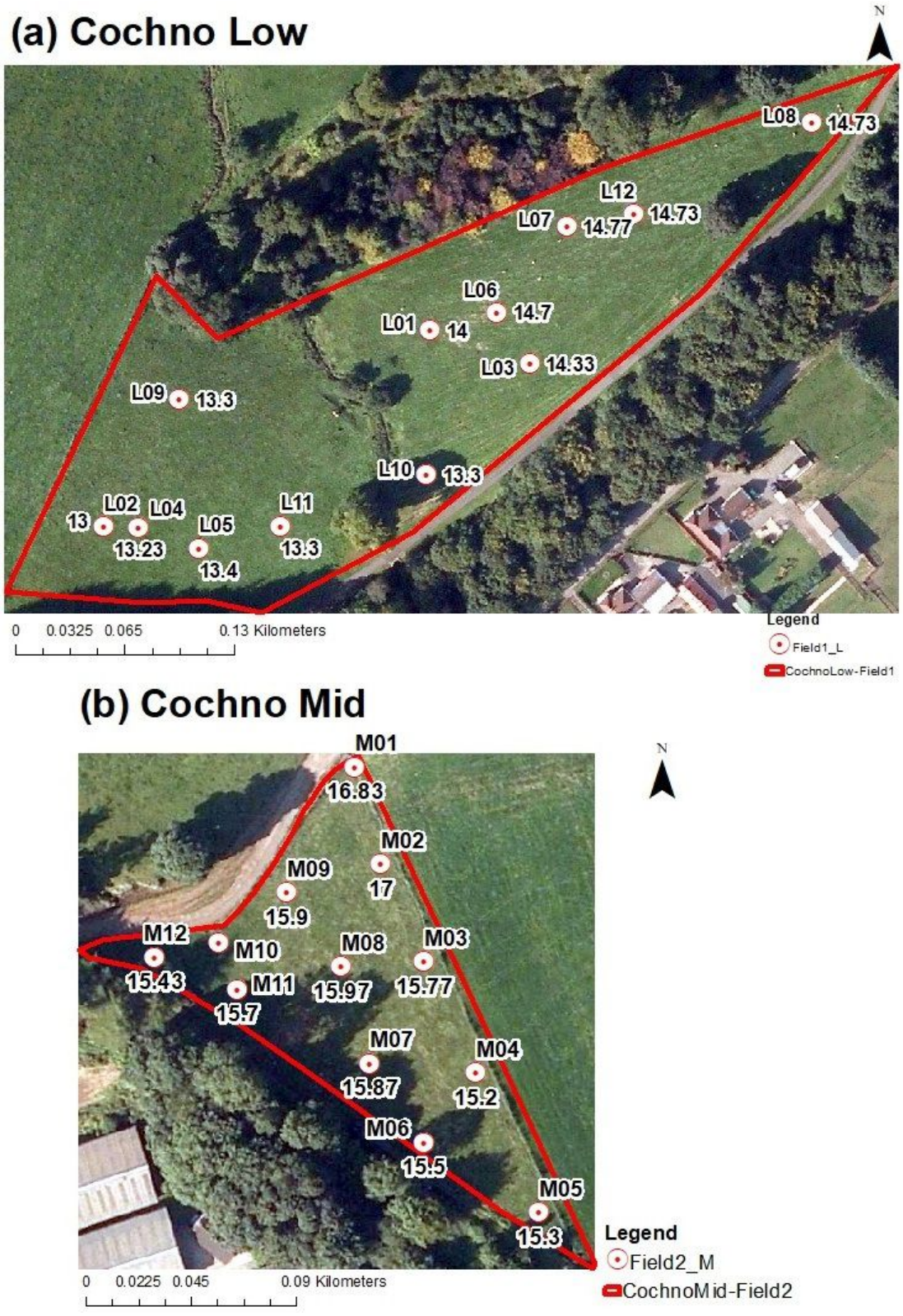
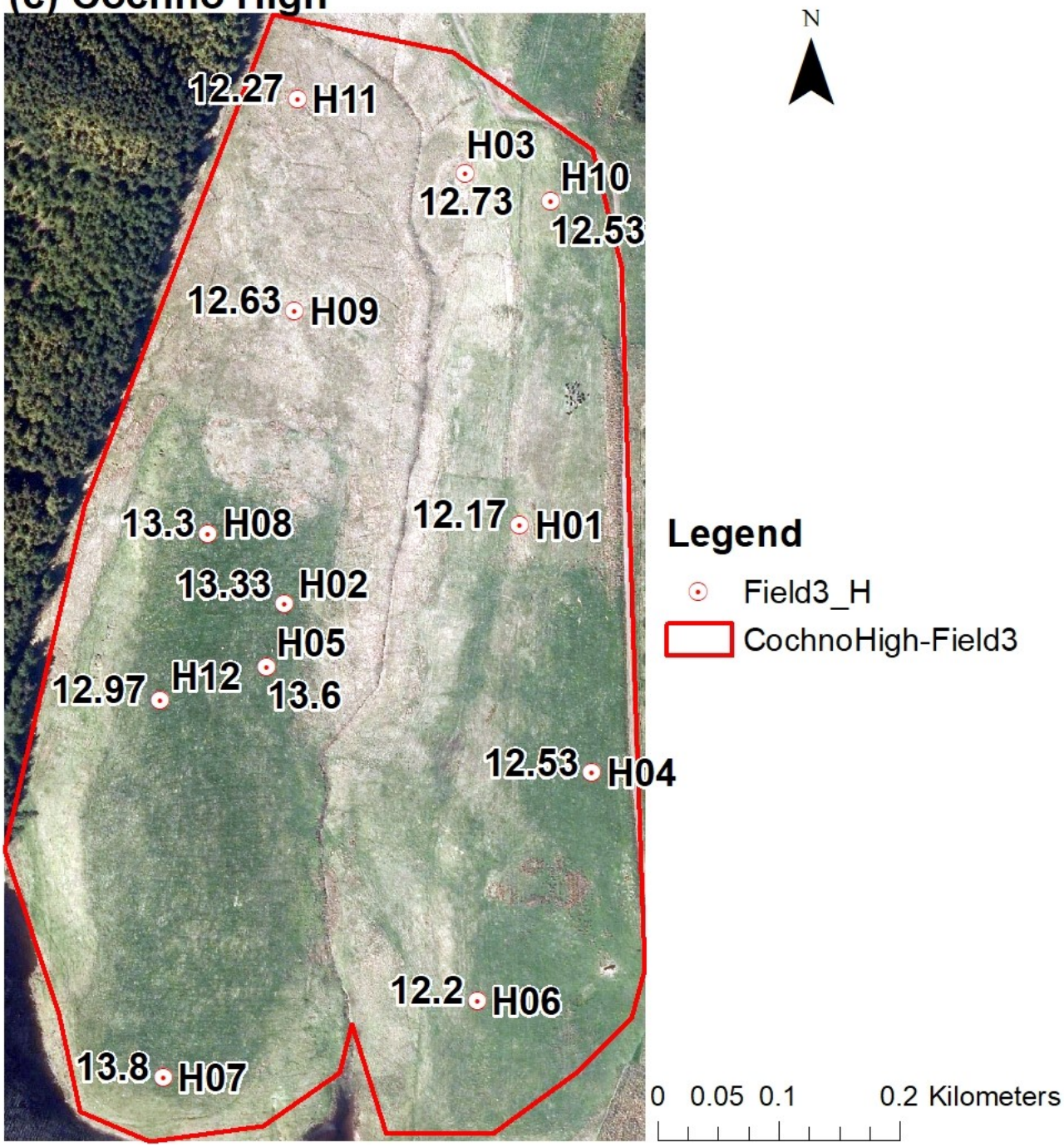
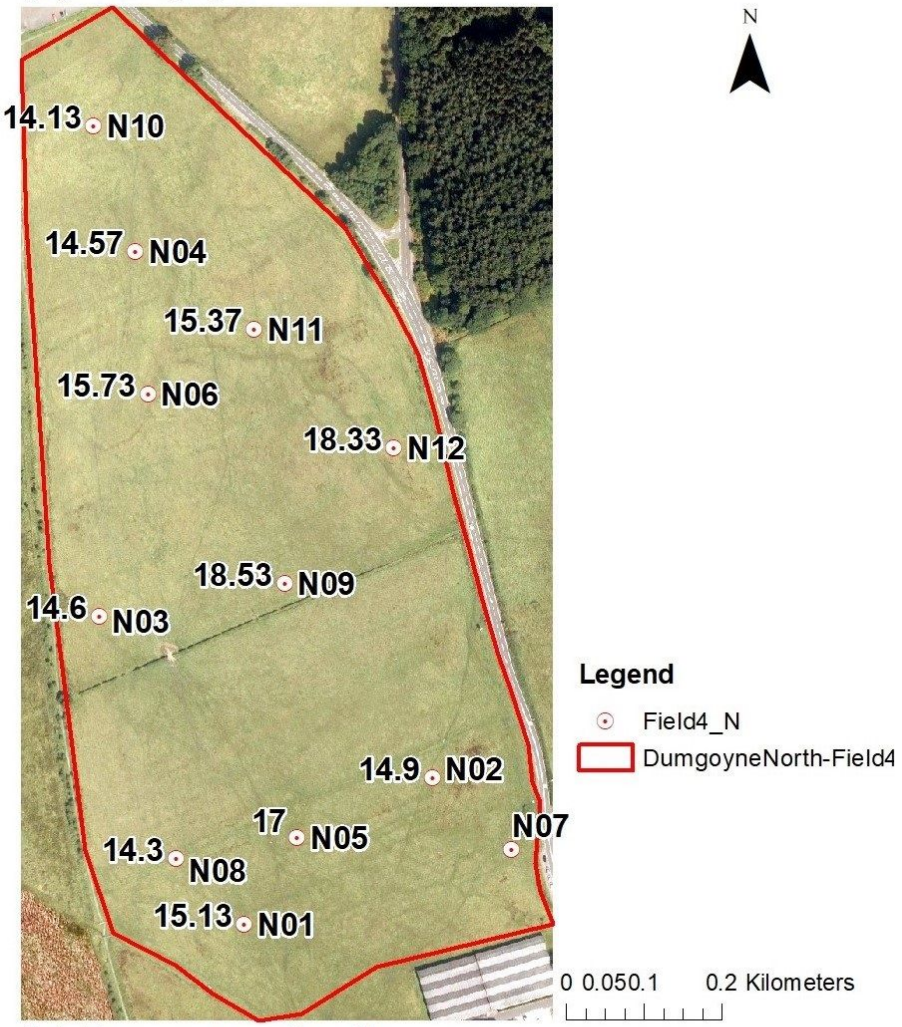


Figure 4.24: The soil temperature (degrees Celsius) for sampling points within the study fields: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South.

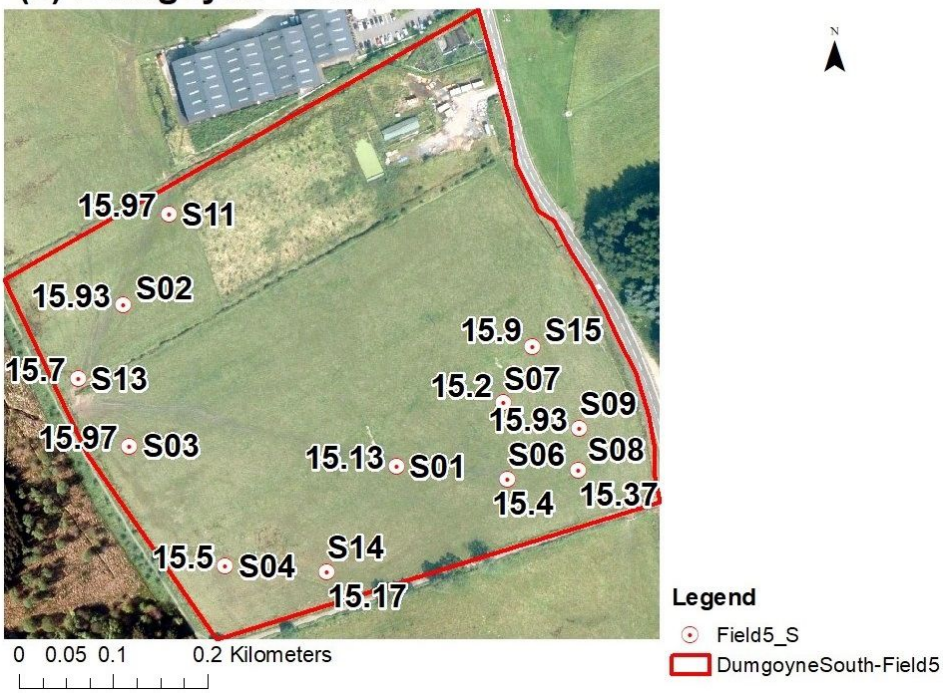
(c) Cochno High



(d) Dumgoyne South



(e) Dumgoyne South



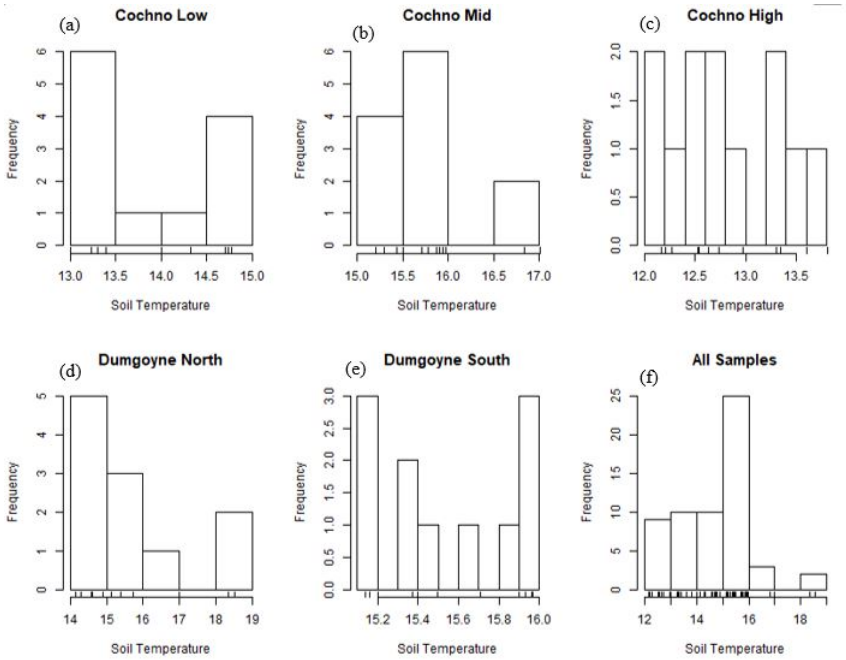


Figure 4.25: Histograms of soil temperature (degrees Celsius) for each field and for all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples.

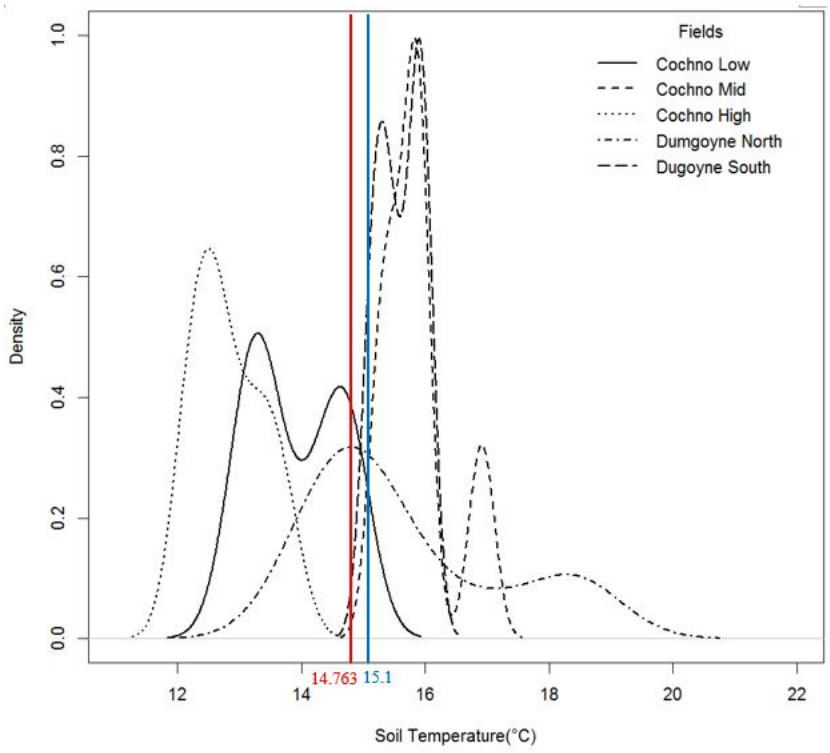


Figure 4.26: Density curve of soil temperature for each field.

4.1.5 Pasture

4.1.5.1 Grass Height

The fields in this study are from typical pasture-based farms, and the risk of infection can be determined by identifying both previous and future use of pastures, particularly the areas that will be grazed by ewes and lamb (NADIS Parasite Report). The grass height was recorded in the field survey at each of the sampling locations, representing the volume of forage grass within each quadrat area (0.25 m^2). The field suvey was conducted on 4 Septemberhe grass height is expected to be variable over time. The average grass height of the five fields was 15.007 cm, and the maximum was 25.33 cm in Dumgoyne North (N03). Figure 4.27 displays the grass height values for each of the sampling points.

Observing the histograms of grass height for each field (Figure 4.28), Dumgoyne North and Cochno Mid were lush with grazing grass, and the grass height of Cochno Mid seemed to be the lowest among the fields. The Shapiro-Wilk normality test applied to the 60 grass heights, showed that the grass height followed a normal distribution. The kernel density curves in Figure 4.29 showed the distribution pattern of grass height and the variation among fields. The red line is the mean of total grass height, and the blue line is the median of total grass height. The density curve of Cochno High and Dumgoyne North were different from the other three fields, appearing on the right of the red line. The distribution of Cochno Mid had the very similar pattern to Dumgoyne South, just farther away from the red line.

The grass height was normally distributed and the homogeneity of variances within each field was passed by Bartlett's test (p-value = 0.9981). The One-way analysis of variance was carried out by R function to test if there was a significant difference among fields. The One-way ANOVA result of grass height was given in Appendix B. The F-test value was 26.44, and the p-value was nearly 2.96E-12. As a result, the null hypothesis cannot be accepted at the 95% confidence interval level. Therefore, we concluded that at least one field was significantly different from the others. South facing areas will receive more sunlight and could have higher grass heights.

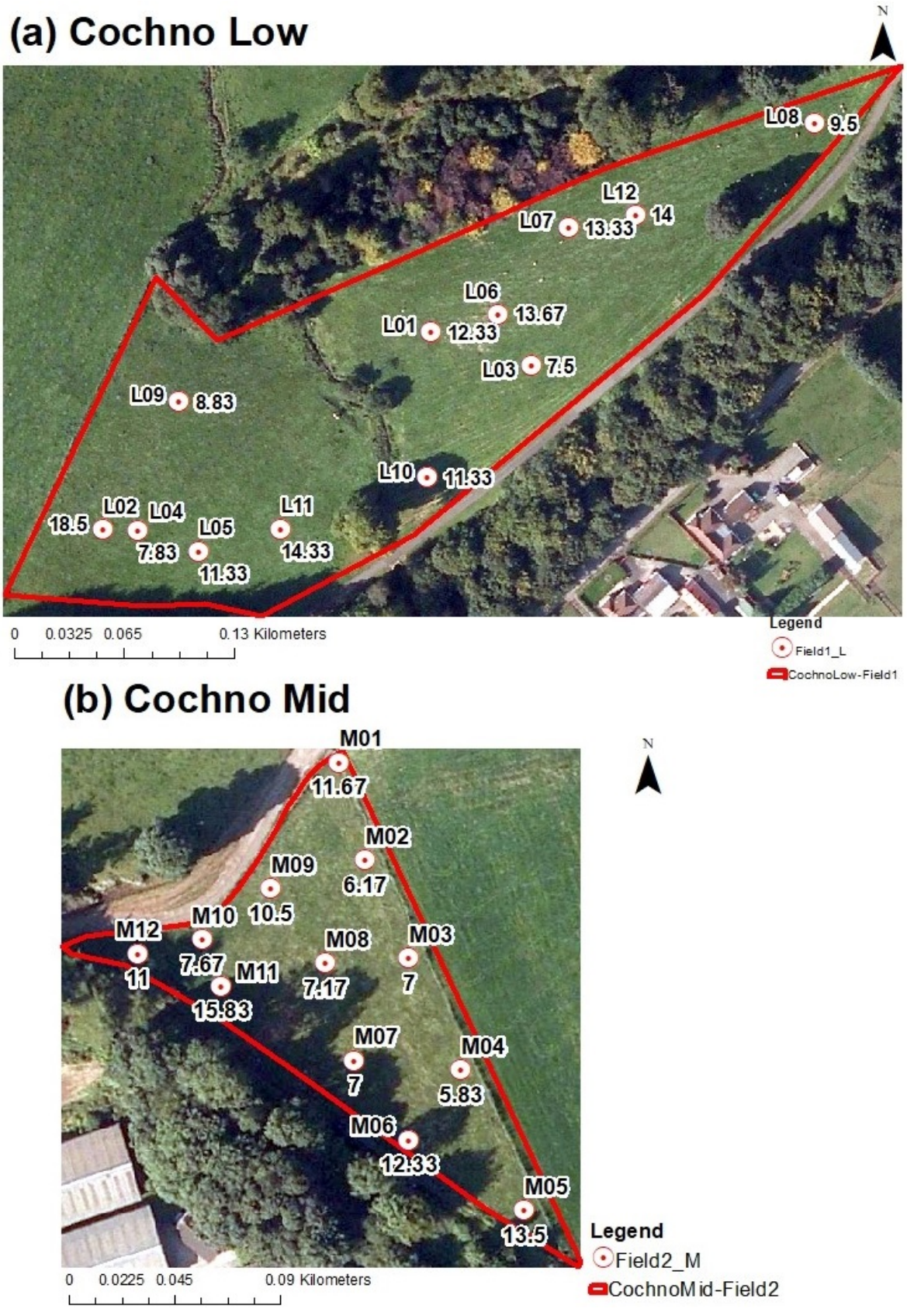
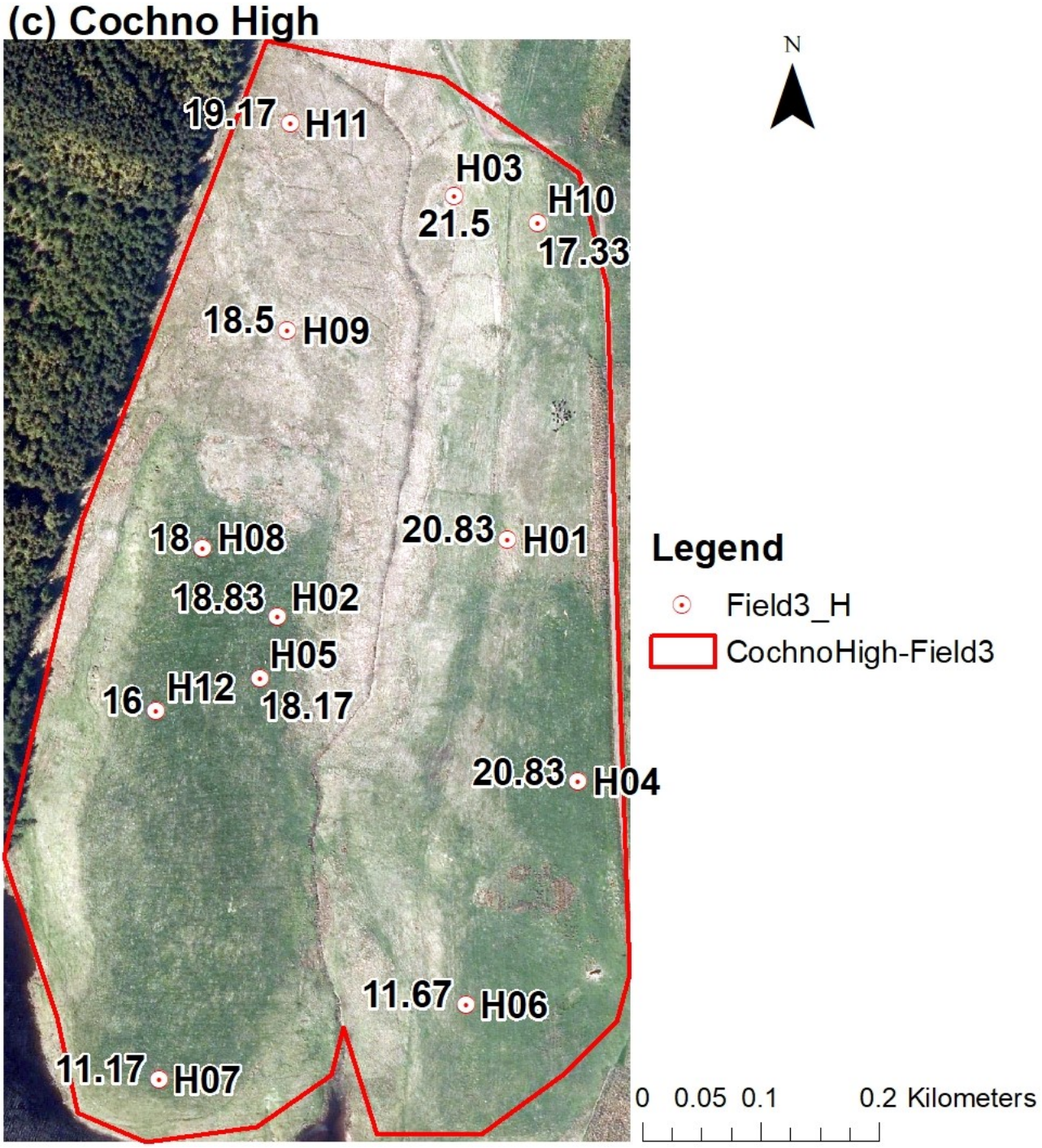
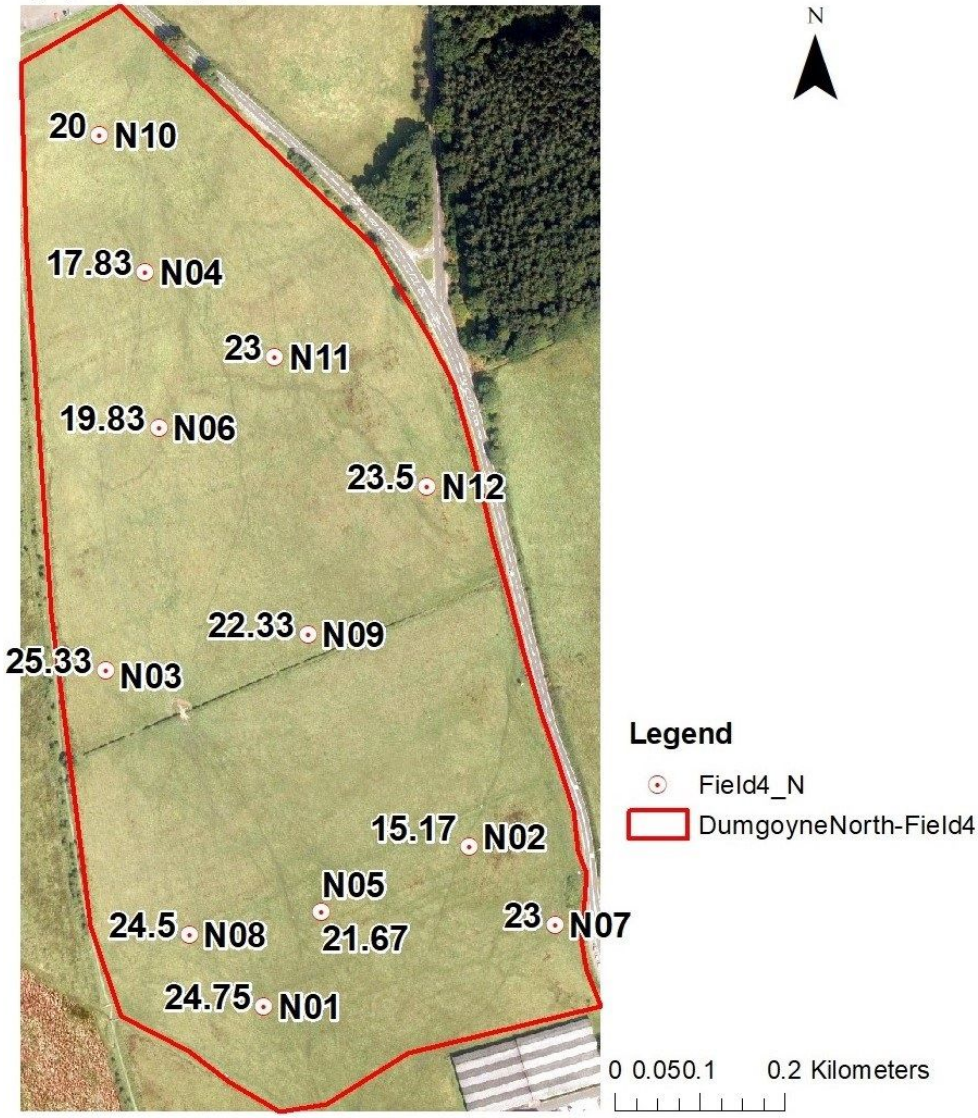


Figure 4.27: The grass height (cm) for sampling points within the study fields: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South.


(c) Cochno High
N
19.17
H11
H03
21.5
H10
17.33
18.5
H09
18
H08
20.83
H01
18.83
H02
H05
16
H12
18.17
20.83
H04
11.67
H06
11.17
H07
Legend
Field3_H
CochnoHigh-Field3
0 0.05 0.1 0.2 Kilometers

(d) Dumgoyne South



(e) Dumgoyne South

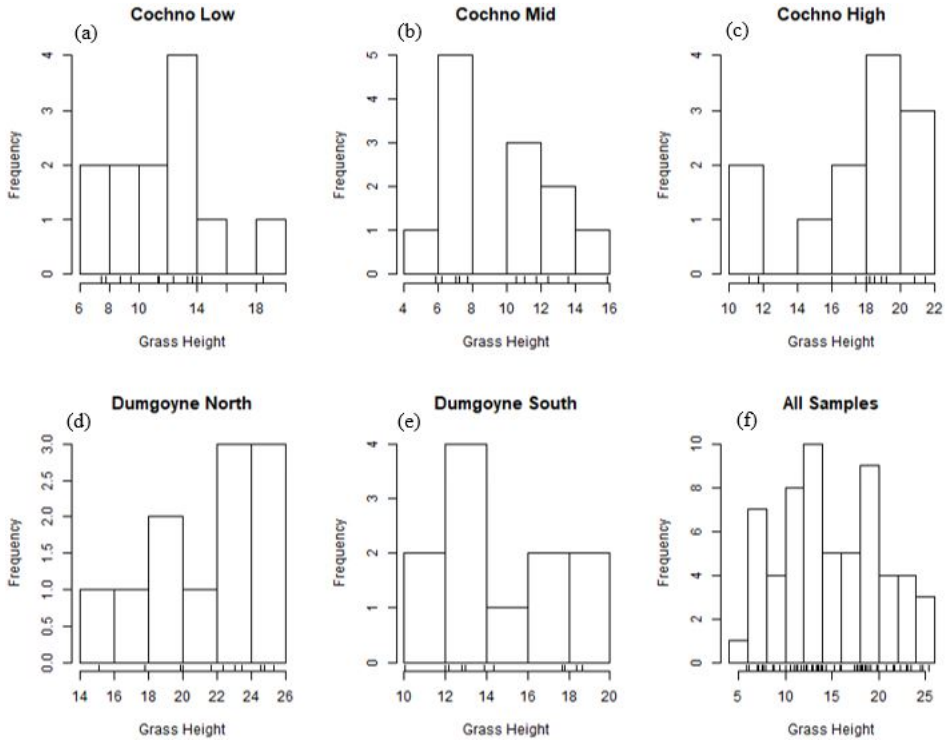


Figure 4.28: Histograms of grass height (cm) for each field and all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples.

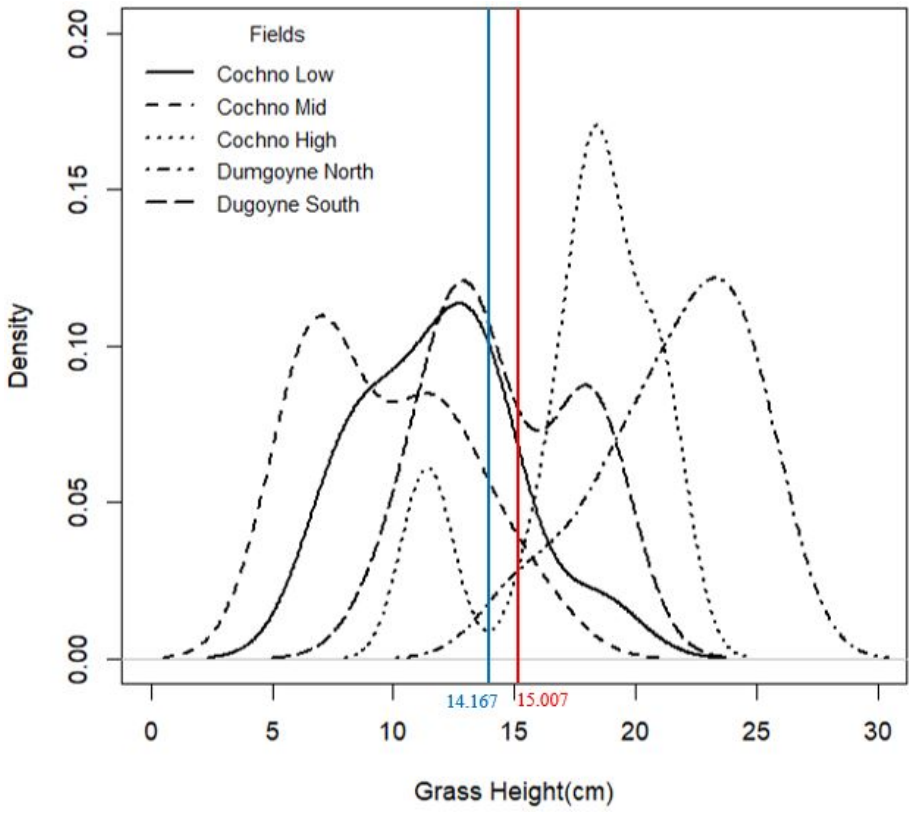


Figure 4.29: Density curve of grass height for each field.

4.1.5.2 Grass Weight

The grass weight was the total wet weight of grass within a $50cm \times 50cm$ quadrat, representing the volume of forage grass and the condition of pasture. The highest grass weight among all samples was H04 in the Cochno High (820 g), and the lowest grass weight was L03 in the Cochno Low (89 g). There was a great disparity between sampling points. The grass weight values for each sampling point for the fields is displayed in Figure 4.30.

According to the Shapiro-Wilk normality test result, the p-value (<0.001) was less than 0.05, and the grass weight of all samples was not normally distributed. The histogram (Figure 4.31) and the kernel density curve (Figure 4.32) of each field below illustrated how the distribution of grass weight varied among fields. In Figure 4.32, the red line is the mean of all samples, and the blue line is the median of all samples. The distribution of grass weight data for most fields were highly skewed to right, apart from the data in Dumgoyne North which was skewed to left and had a higher density on the right side of the red line. As with the grass heights, Cochno Mid and Dumgoyne South were similar in the density curves of the grass weight. The grass weight did not follow the normal distribution and could not be transformed into the normal distribution. Hence, the non-parametric test, Kruskal-Wallis test, was conducted to analyse the variance. The test result is in Appendix B, and the p-value is nearly zero (3.246e-05). This revealed that the average grass weight of each field was not significantly equal and the variation of grass weight by fields was significant.

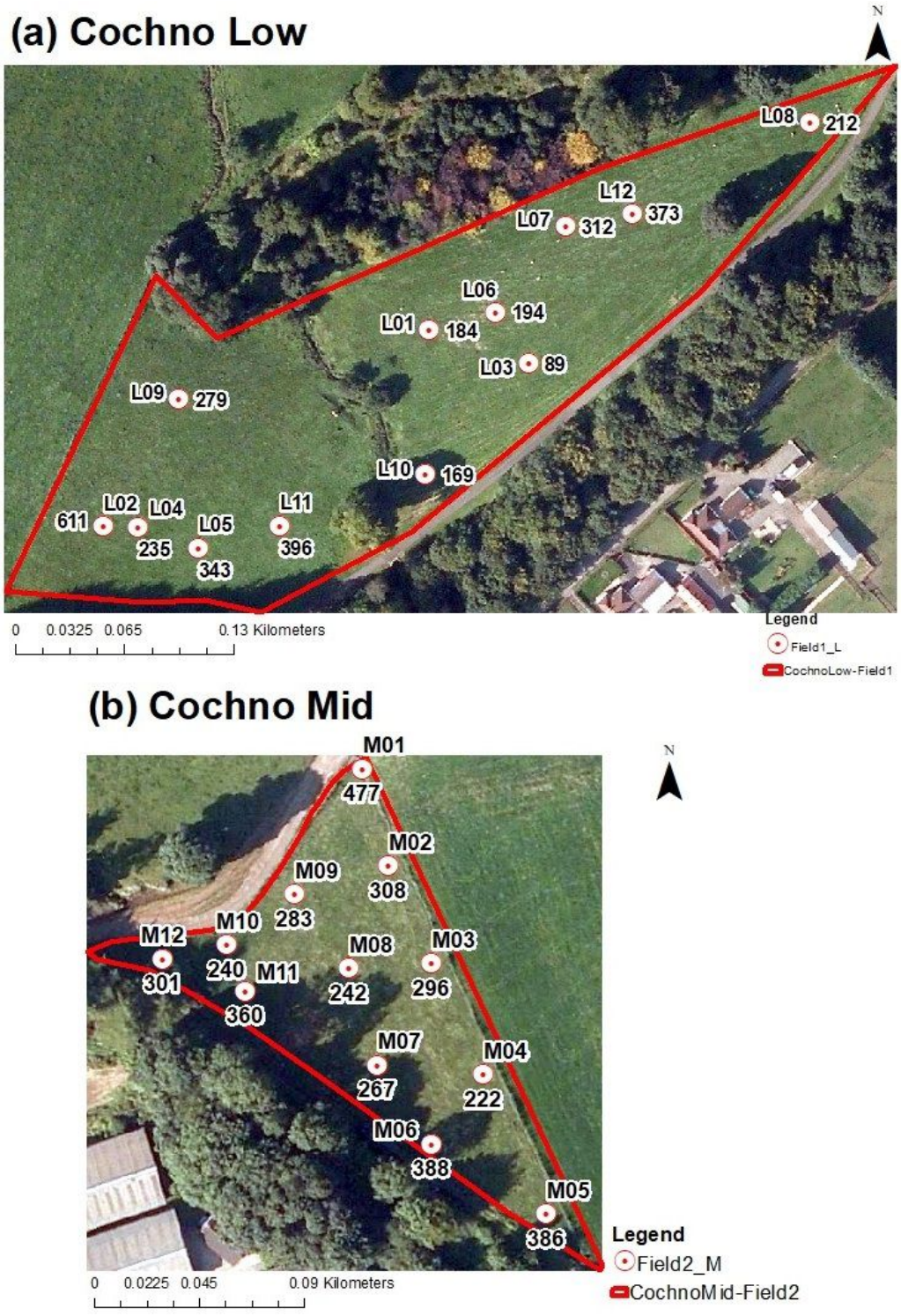
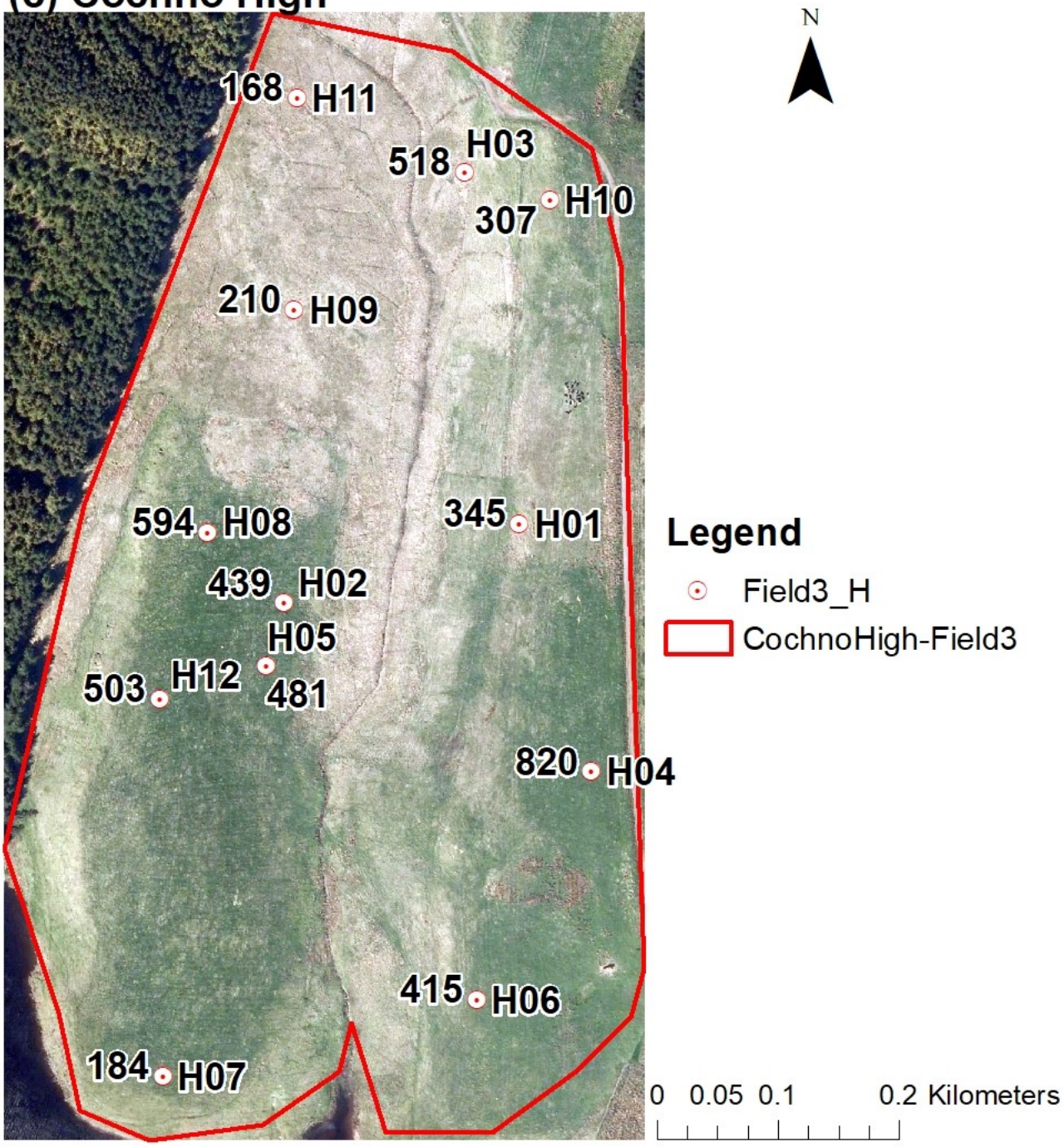
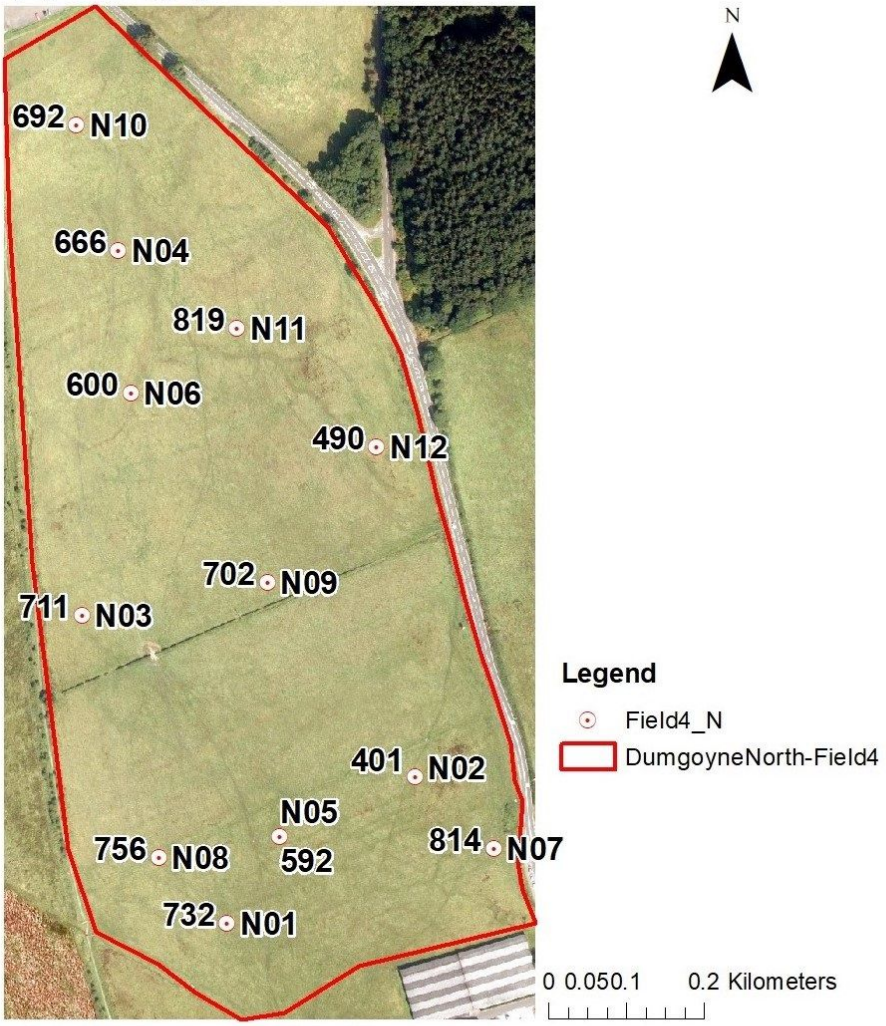


Figure 4.30: The grass weight values for sampling points within the study fields: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South.

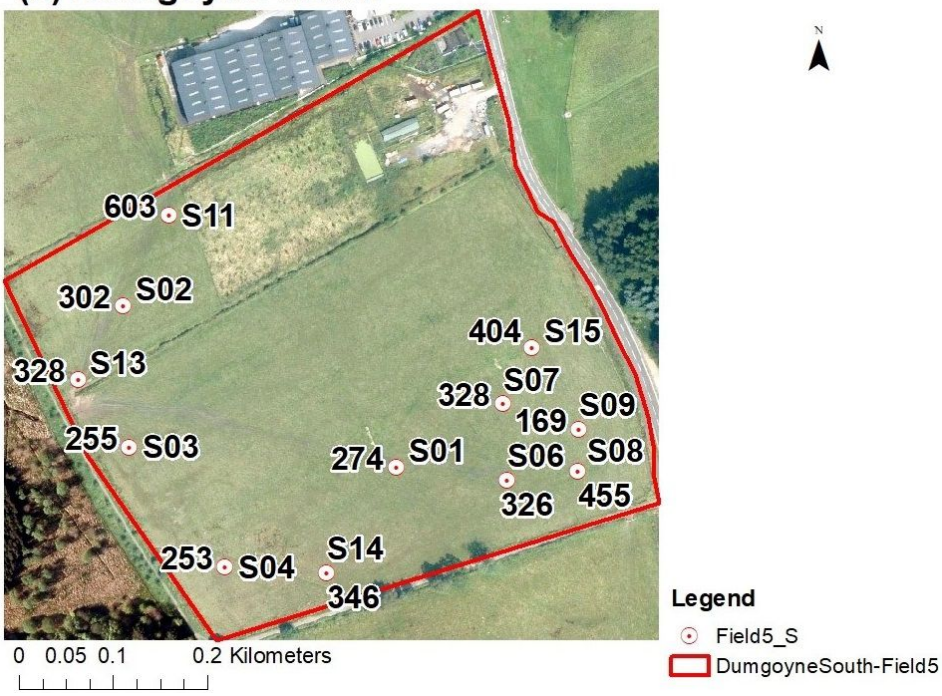
(c) Cochno High



(d) Dumgoyne South



(e) Dumgoyne South



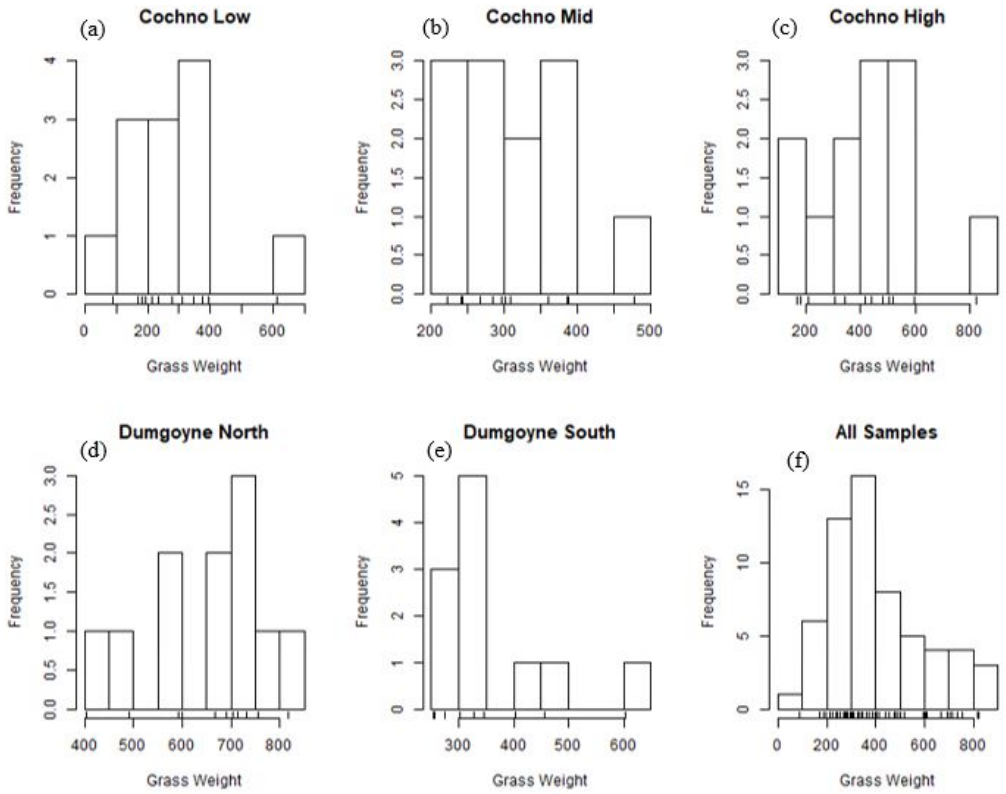


Figure 4.31: Histograms of grass height for each field and all samples: (a) Cochno Low; (b) Cochno Mid; (c) Cochno High; (d) Dumgoyne North; (e) Dumgoyne South; (f) All samples.

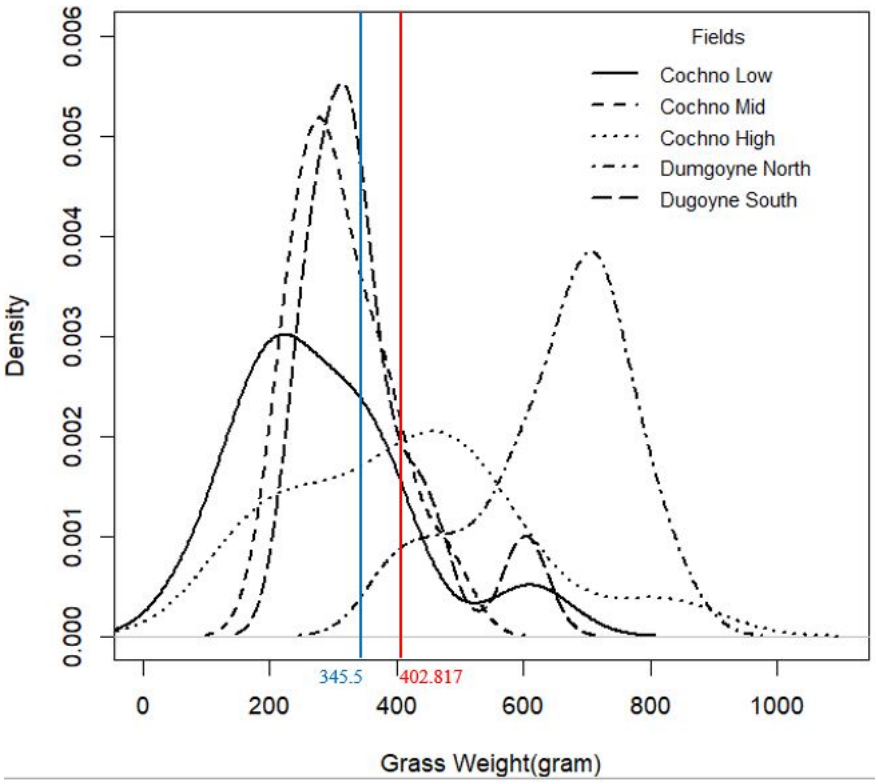


Figure 4.32: Density curve of grass weight for each field.

4.2 Relationships between remote sensing, terrain, soil and pasture datasets

Initial assessment of relationships among variables within each study area was carried out using statistical tests. The strength of relationships was assessed by Pearson product moment correlation, and the correlation coefficients (r) of each field represent the strength and direction of a linear relationship between two variables.

4.2.1 Cochno Low

Pearson's correlation coefficients matrix of all variables for the field Cochno Low are listed in Table 4.1 and reveal significant (at 95% confidence level) correlations among variables. There were 12 samples and 11 variables in this analysis. Yield and count of metacercariae are two critical variables indicting *F. hepatica* level, and their univariate association with environmental variables can provide a preliminary indication of how the environment might affect the density of *F. hepatica*. In Cochno Low, Soil Moisture was the only variable that showed a strong positive association (r = 0.591) with Yield, whereas the coefficient between Soil Moisture and metacercariae was 0.44. There were also quite strong positive associations be- tween remote sensing index variables (NDWI and NDVI) and *F. hepatica* level (Yield and Metacercarial cyst counts), but there were also negative associations. What was unexpected was that all of the topographical factors showed very weak associations with both Yield and Metacercarial cyst counts.

The relationships among environmental factors is worth noting due to the strong relationships between variables cause multicollinearity for predictive models. For example, the correlation matrix reflected a powerful positive linear association between NDVI and NDWI. Moreover, Soil Temperature also showed strong negative relationships with both NDWI and NDVI. There was also a similar relationship between Soil Moisture and Elevation. Furthermore, there were some unexpected connections revealed in the Pearson's correlation coefficient matrix. For instance, the NDVI, which was the way to measure healthy green vegetation, showed a weak negative relationship with the Grass Height, the real measurement from field survey. Similarly, the NDWI, as an index to indicate land surface moisture, had nearly no linear relationship with Soil Moisture.

Table 4.1: Pearson's Correlation Coefficient of Cochno Low

	Yield	Metacercaria	GrassHeight	GrassWeight	SoilMoisture	SoilTemperature	NDWI	NDVI	Elevation	Slope	Aspect
Yield	1										
Metacercaria	0.86	1									
GrassHeight	-0.34	-0.19	1								
GrassWeight	-0.60	-0.46	0.78	1							
SoilMoisture	0.59	0.44	-0.20	-0.36	1						
SoilTemperature	0.30	0.27	-0.07	-0.38	-0.36	1					
NDWI	-0.61	-0.58	-0.14	0.40	-0.08	-0.70	1				
NDVI	-0.54	-0.59	-0.18	0.34	-0.07	-0.64	0.96	1			
Elevation	-0.27	-0.22	-0.01	0.35	-0.51	0.26	0.28	0.23	1		
Slope	-0.21	-0.13	-0.15	0.30	-0.25	-0.23	0.46	0.34	0.62	1	
Aspect	0.19	-0.10	-0.02	-0.36	-0.09	0.32	-0.54	-0.44	-0.42	-0.49	1

4.2.2 Cochno Mid

Table 4.2 shows the Pearson's correlation coefficients for Cochno Mid. There were 12 samples and 11 variables used in the analysis. Firstly, there were strong negative linear relationships between NDVI and *F. hepatica* (Yield and Metacercarial cysts). Despite Grass Weight negatively relating to Yield, other variables did not show linear relationships with *F. hepatica*. Three pairs of independent variables showed strong positive associations: Grass Weight with Slope, Soil Temperature with Elevation, and NDWI with Elevation. There were also three pairs of negatively associated variables: NDWI with Grass Height, Soil Moisture with Slope, and Soil Moisture with Aspect. In contrast to Cochno Low, there were few strong associations among the variables in Cochno Mid.

Table 4.2: Pearson's Correlation Coefficient of Cochno Mid

	Yield	Metacercaria	GrassHeight	GrassWeight	SoilTemperature	SoilMoisture	NDWI	NDVI	Elevation	Slope	Aspect
Yield	1										
Metacercaria	0.99	1									
GrassHeight	-0.20	-0.12	1								
GrassWeight	-0.32	-0.21	0.71	1							
SoilTemperature	0.02	0.02	-0.22	0.29	1						
SoilMoisture	-0.20	-0.16	-0.12	0.03	0.02	1					
NDWI	-0.16	-0.26	-0.61	-0.34	0.45	0.04	1				
NDVI	-0.59	-0.63	-0.28	-0.19	-0.03	0.39	0.41	1			
Elevation	0.07	0.04	-0.10	0.19	0.81	-0.39	0.50	-0.26	1		
Slope	-0.26	-0.24	0.20	0.57	0.34	-0.53	0.11	-0.23	0.57	1	
Aspect	0.11	0.01	0.004	-0.22	0.23	-0.57	0.37	-0.003	0.54	0.15	1

4.2.3 Cochno High

The relationships among variables in Cochno High are shown in Table 4.3.4.3.There were 12 samples and 11 variables in Cochno High. The correlations between environmental factors and were limited, as with the other two fields at Cochno. Soil Moisture was the only variable that had strong positive associations with both Yield and Metacercarial cysts; otherwise, none of

the variables had significant negative relationships with *F. hepatica*, and Soil Temperature was moderately associated with Yield and Metacercarial cysts. Slope showed a strong positive association with Metacercarial cysts but was not significantly associated with Yield.

Focusing on relationships among environmental factors at Cochno High, Grass Height and NDVI had strong positive correlations with Elevation. Soil Moisture was significantly positively related to topographical factors, Slope and Aspect (r = 0.807 and r = 0.701 respectively). Moreover, Soil Temperature also had strong negative relationships with both Slope and Aspect. Just like the other two fields at Cochno, more than half of the variables showed very weak correlations. The correlation coefficient of NDVI and Grass Weight was very low (r = 0.001), which meant that there was no linear correlation between NDVI and Grass Weight. Also, NDWI and Soil Moisture were only very weakly correlated.

Table 4.3: Pearson's Correlation Coefficient of Cochno High

	Yield	Metacercaria	GrassHeight	GrassWeight	SoilTemperature	SoilMoisture	NDWI	NDVI	Elevation	Slope	Aspect
Yield	1										
Metacercaria	0.98	1									
GrassHeight	0.38	0.38	1								
GrassWeight	-0.20	-0.07	0.35	1							
SoilTemperature	-0.40	-0.31	-0.29	0.05	1						
SoilMoisture	0.73	0.71	0.22	-0.33	-0.30	1					
NDWI	-0.12	-0.18	0.45	-0.24	0.08	-0.15	1				
NDVI	0.13	0.14	0.56	0.001	0.13	0.01	0.82	1			
Elevation	0.34	0.31	0.59	0.07	-0.60	0.48	0.34	0.44	1		
Slope	0.49	0.53	-0.06	-0.22	-0.15	0.81	-0.16	0.06	0.39	1	
Aspect	0.33	0.34	0.16	-0.16	-0.50	0.70	-0.11	-0.18	0.49	0.59	1

4.2.4 Dumgoyne North

Pearson's correlation coefficients of all variables for the field Dumgoyne North are given listed in Table 4.4 and reveal significant (95% confidence level) correlations among variables. There were 11 samples and 11 variables in Dumgoyne North used in this analysis, due to the missed sample N07. Grass Weight was the only independent variable significantly correlated with *F. hepatica* at Dumgoyne North, which had a strong positive relationship with both Yield and Metacercarial cysts. Remote sensing indices (NDWI and NDVI) had moderate positive relationships with the density of *F. hepatica* , and Soil Moisture had a weak negative association with *F. hepatica*. Soil Moisture was significantly negatively associated (r = -0.828) with NDWI and had a strong negative relationship with NDVI as well. Slope had strong negative relationships with NDWI.

Table 4.4: Pearson's Correlation Coefficient of Dumgoyne North

	Yield	Metacercaria	GrassHeight	GrassWeight	SoilTemperature	SoilMoisture	NDWI	NDVI	Elevation	Slope	Aspect
Yield	1										
Metacercaria	1.00	1									
GrassHeight	0.14	0.14	1								
GrassWeight	0.50	0.51	0.60	1							
SoilTemperature	-0.07	-0.08	0.17	-0.29	1						
SoilMoisture	-0.40	-0.40	-0.10	-0.05	-0.40	1					
NDWI	0.36	0.36	-0.02	-0.15	0.46	-0.83	1				
NDVI	0.34	0.35	0.03	0.24	0.12	-0.61	0.74	1			
Elevation	0.12	0.12	-0.30	-0.39	0.49	-0.48	0.42	0.26	1		
Slope	0.11	0.12	0.45	0.34	-0.34	0.36	-0.54	-0.46	-0.36	1	
Aspect	-0.13	-0.12	0.56	-0.15	0.39	-0.02	0.13	0.06	0.09	0.16	1

4.2.5 Dumgoyne South

Correlations among variables in Dumgoyne South differed from those of Dumgoyne North. There were 11 samples and 11 variables in Dumgoyne South. The Table 4.5 4.5 lists correlations among variables within Dumgoyne South. Soil Moisture had strong positive associations with both Yield and Metacercarial cysts, whereas Soil Temperature and *F. hepatica* had no significant associations. Grazing factors (Grass Height and Grass Weight) had negative associations with *F. hepatica*. Soil Temperature had significant positive correlations with remote sensing indexes (NDWI and NDVI). There were strong positive relationships between Slope and grazing factors (Grass Height and Grass Weight).

Table 4.5: Pearson's Correlation Coefficient of Dumgoyne South

	Yield	Metacercaria	GrassHeight	GrassWeight	SoilTemperature	SoilMoisture	NDWI	NDVI	Elevation	Slope	Aspect
Yield	1										
Metacercaria	0.98	1									
GrassHeight	-0.55	-0.51	1								
GrassWeight	-0.53	-0.47	0.65	1							
SoilTemperature	0.06	0.08	-0.32	0.27	1						
SoilMoisture	0.57	0.63	-0.11	-0.23	0.11	1					
NDWI	-0.24	-0.23	-0.26	0.35	0.83	-0.22	1				
NDVI	-0.45	-0.45	-0.03	0.42	0.56	-0.49	0.87	1			
Elevation	-0.59	-0.63	0.36	0.30	-0.32	-0.79	-0.01	0.12	1		
Slope	-0.52	-0.41	0.58	0.72	0.30	0.28	0.26	0.19	-0.06	1	
Aspect	-0.08	-0.003	0.20	0.17	-0.21	-0.03	-0.11	0.12	0.03	0.16	1

Chapter 5

Relationships between *F. hepatica* and Environmental Factors

5.1 Univariate regression models

Including all environmental factors in the regression leads to the problem of multicollinearity. Several of the environmental factors in this study demonstrated a high degree of correlation in Chapter 3.2. To reduce the potential for multicollinearity among environmental factors, univariate negative binomial regression was conducted on each independent variable first and only the variables with significant coefficients were selected for use in the multivariate regression model. However, the overdispersion of the model has impacts on the model accuracy. The goodness of fit of each univariate regression model was analysed using the chi-square goodness of fit test. McFadden's Pseudo R^2 and AIC values were also used to test the goodness of fit of the regression model.

5.1.1 Modeling remote sensing indices against *F. hepatica* density in pasture

The Yield2 data were modelled against the remote sensing indices across the study fields, normalized difference water index (NDWI) and normalized difference vegetation index (NDVI) individually as the nested factors within fields. Models assumed a negative binomial distribution of the response, where a generalised linear model with a logit link function was specified. Cochno High was used as the reference group, and the other four fields were the comparison groups. The following models were examined:

- *Yield2* $\sim$ *Field * NDWI*
- *Yield2* $\sim$ *Field * NDVI*

There were no consistent effects of field, NDWI, NDVI or their interactions on the density of *F. hepatica* Neither field nor NDWI nor their interaction were significantly associated with Yield2. The regression results of NDWI model are shown in Table 5.1, and the fitted line plot shows the regression results graphically (Figure 5.1). Although the effects of Cochno Mid (increased Yield2) and its interaction with NDVI (in Cochno Mid, increased NDVI reduced Yield2) were significant (Table 5.2), it is evident from Figure 5.2 that the effects are small and do not represent a consistent pattern. The Chi-squared test of each covariate was carried out to examine the goodness of fit for each model (Table 5.3). The NDVI model fitted better than the NDWI model, with more significant factors and reduced residual deviance.

Table 5.1: Univariate regression model of NDWI

	Estimate	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept)	-3.8680	1.9816	-1.9519	0.0568
CochnoLow	2.4184	2.5319	0.9552	0.3443
CochnoMid	1.6021	2.7452	0.5836	0.5622
DumgoyneNorth	-27.8168	24.2147	-1.1488	0.2563
DumgoyneSouth	1.6155	5.1755	0.3121	0.7563
NDWI	-2.4234	5.6189	-0.4313	0.6682
CochnoLow:NDWI	-6.4719	7.6604	-0.8449	0.4024
CochnoMid:NDWI	-1.4933	7.5295	-0.1983	0.8436
DumgoyneNorth:NDWI	69.3453	59.9827	1.1561	0.2534
DumgoyneSouth:NDWI	-4.9382	15.6377	-0.3158	0.7535

Table 5.2: Univariate regression model of NDVI

	Estimate	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept)	-5.7334	1.9127	-2.9976	0.0043
CochnoLow	4.1896	2.4138	1.7357	0.0890
CochnoMid	12.4122	3.7541	3.3063	0.0018
DumgoyneNorth	-73.1315	60.6895	-1.2050	0.2341
DumgoyneSouth	14.0834	11.5870	1.2155	0.2301
NDVI	2.4102	4.3133	0.5588	0.5789
CochnoLow:NDVI	-6.4816	4.8737	-1.3299	0.1898
CochnoMid:NDVI	-21.4763	7.4540	-2.8812	0.0059
DumgoyneNorth:NDVI	86.5671	72.8818	1.1878	0.2408
DumgoyneSouth:NDVI	-19.7320	15.9123	-1.2400	0.2210

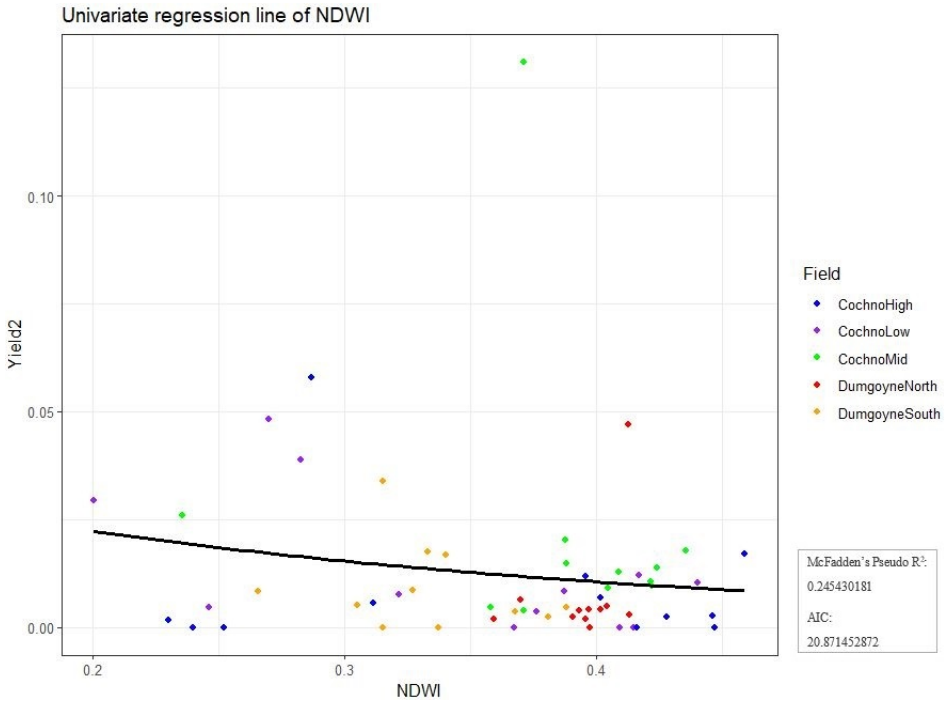


Figure 5.1: The dot plot and fitted line of NDWI against Yield2.

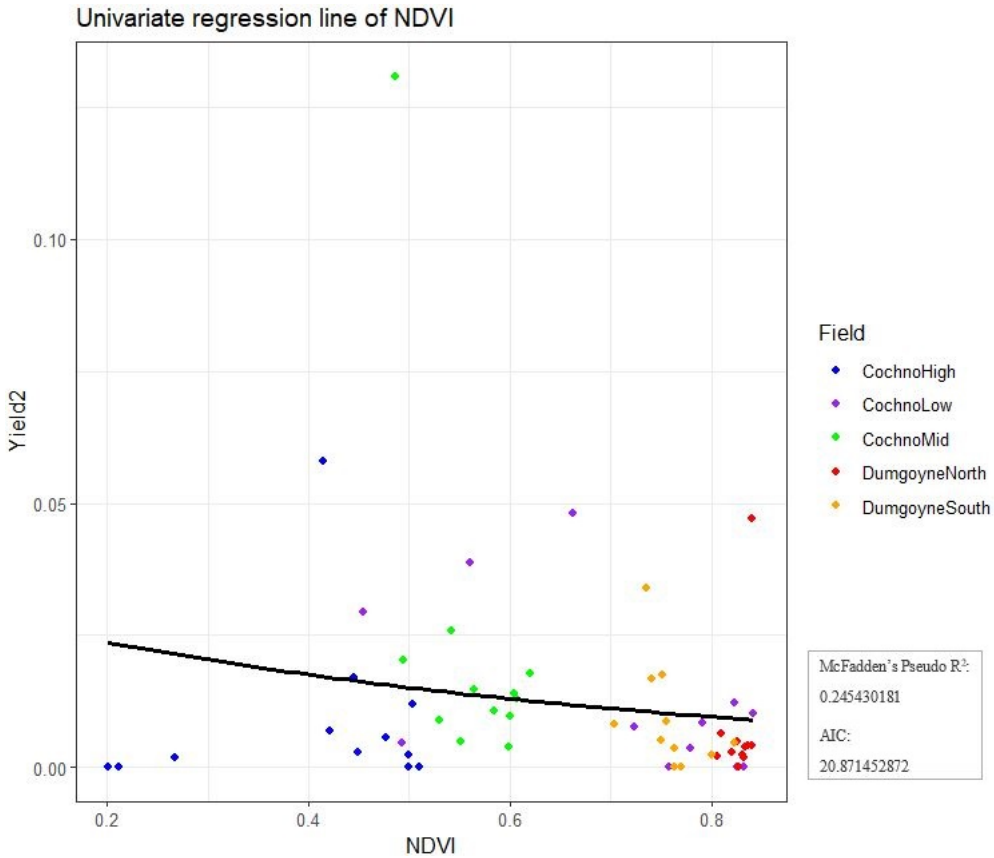


Figure 5.2: The dot plot and fitted line of NDVI against Yield2.

Table 5.3: Chi-squared goodness of fit for remote sensing index models

Models	Variables	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
Yield2~Field *NDWI	NULL			57	1.1549	
	Field	4	0.1393	53	1.0156	0.2293
	NDWI	1	0.0705	52	0.9451	0.0918
	Field:NDWI	4	0.0736	48	0.8715	0.5630
Yield2~Field * NDVI	NULL			57	1.1549	
	Field	4	0.1393	53	1.0156	0.0942
	NDVI	1	0.0887	52	0.9268	0.0247
	Field:NDVI	4	0.2245	48	0.7024	0.0124

5.1.2 Modeling Soil properties against *F. hepatica* density in pasture

As with the analysis on *F. hepatica* level, the soil properties were also appropriate to model across the fields. Soil properties, such as temperature and moisture were used as nested variables within fields to perform a univariate negative binomial regression model against Yield2 separately. The two models were:

- *Yield2* $\sim$ *Field * SoilMoisture*
- *Yield2* $\sim$ *Field * SoilTemperature*

When modeling Soil Moisture to *F. hepatica* level, Cochno Mid, Dumgoyne North, Soil Moisture and Intercept had p-values less than 0.05 (Table 5.4). The effect estimates of Cochno Mid, Dumgoyne North and Soil Moisture were 8.479, 12.559 and 9.434, indicating that they were all positively related to Yield2. The fitted regression line for Soil Moisture against Yield2 was shown in the Figure 5.3. It indicates that the significant interaction term of Cochno-Mid:SoilMoisture is likely due to high leverage caused by a single factor. In contrast to Soil Moisture, Soil Temperature did not show any significant effect on *F. hepatica* level in any of the fields. The regression table of Soil Temperature model is shown in Table 5.5 and the fitted regression line is shown in Figure 5.4.

Table 5.4: Univariate regression model of Soil Moisture

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-9.9984	2.1579	-4.6334	0.00003
CochnoLow	-3.9034	5.5482	-0.7035	0.4851
CochnoMid	8.4791	2.8676	2.9569	0.0048
DumgoyneNorth	12.1559	5.3362	2.2780	0.0272
DumgoyneSouth	-1.2376	5.1773	-0.2390	0.8121
SoilMoisture	9.4340	3.4305	2.7500	0.0084
CochnoLow:SoilMoisture	5.6757	8.4935	0.6682	0.5072
CochnoMid:SoilMoisture	-13.0877	4.6311	-2.8260	0.0068
DumgoyneNorth:SoilMoisture	-22.1570	9.7034	-2.2834	0.0269
DumgoyneSouth:SoilMoisture	0.8565	7.9195	0.1082	0.9143

Table 5.5: Univariate regression model of Soil Temperature

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	18.8448	18.0639	1.0432	0.3021
CochnoLow	-30.3465	20.0873	-1.5107	0.1374
CochnoMid	-23.2874	20.4045	-1.1413	0.2594
DumgoyneNorth	-22.3719	19.2124	-1.1645	0.2500
DumgoyneSouth	-26.3787	31.1377	-0.8472	0.4011
SoilTemperature	-1.8662	1.4487	-1.2882	0.2039
CochnoLow:SoilTemperature	2.3811	1.5762	1.5106	0.1374
CochnoMid:SoilTemperature	1.9095	1.5670	1.2186	0.2290
DumgoyneNorth:SoilTemperature	1.7768	1.5084	1.1779	0.2446
DumgoyneSouth:SoilTemperature	2.0493	2.1785	0.9407	0.3516

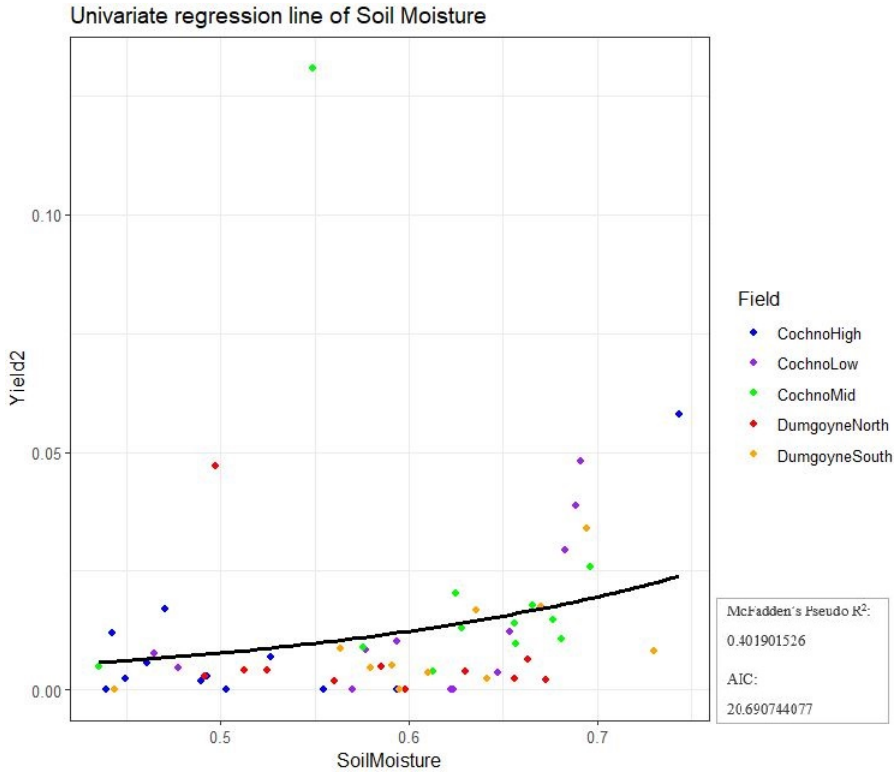


Figure 5.3: The dot plot and fitted line of Soil Moisture against Yield2.

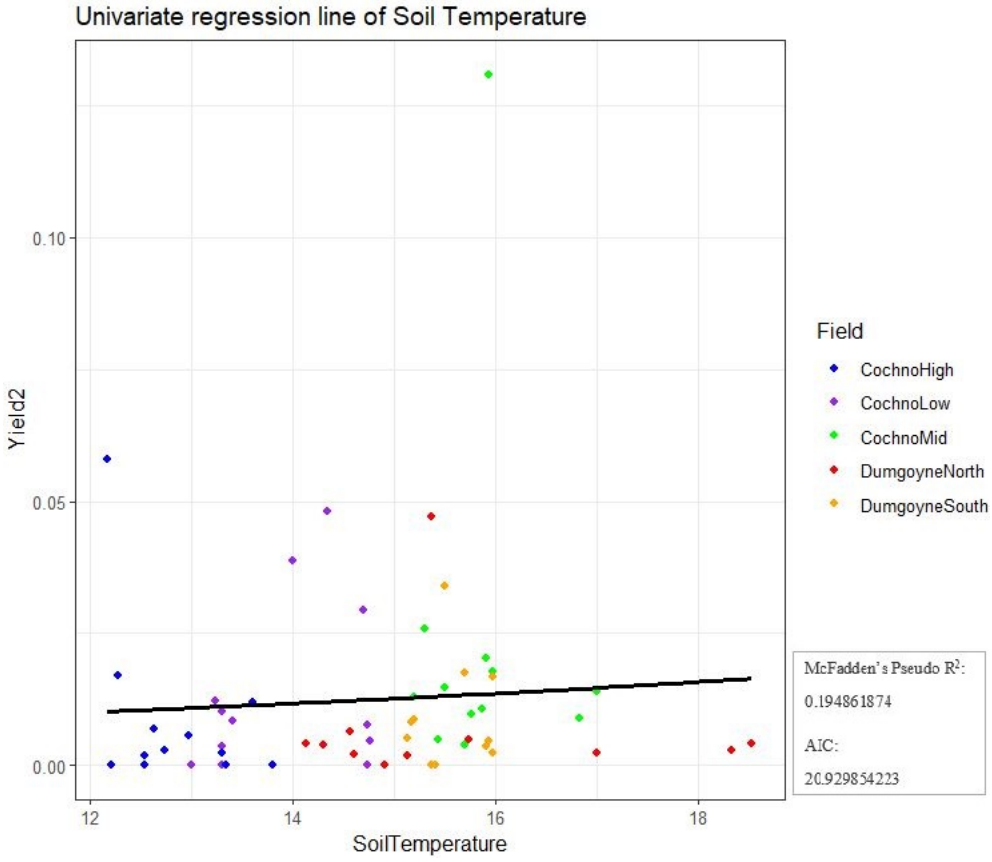


Figure 5.4: The dot plot and fitted line of Soil Temperature against Yield2.

Table 5.6: Chi-squared goodness of fit for soil models

Models	Variables	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
Yield2~Field*SoilMositure	NULL			57	1.1549	
	Field	4	0.1393	53	1.0156	0.0922
	SoilMoisture	1	0.0561	52	0.9595	0.0730
	Field:SoilMoisture	4	0.2687	48	0.6907	0.0039
Yield2~Field*SoilTemperature	NULL			57	1.1549	
	Field	4	0.1393	53	1.0156	0.2777
	SoilTemperature	1	0.0004	52	1.0152	0.9072
	Field:SoilTemperature	4	0.0853	48	0.9299	0.5379

5.1.3 Modeling topographical factors against *F. hepatica* density in pasture

Assessing the relationship between the terrain and the *F. hepatica* density on pasture, the Elevation, Slope, and Aspect were used as the nested covariate to build models for the *F. hepatica* density. Three univariate regression models were examined:

- *Yield2* $\sim$ *Field * Elevation*
- *Yield2* $\sim$ *Field * Slope*
- *Yield2* $\sim$ *Field * Aspect*

Table 5.7 presents the negative binomial regression results of Elevation model and Figure 5.5 illustrates the relationships between Yield2 and Elevation for each field with dot plots. Elevation had no significant association with the Yield2 in any of the fields. Slope tended to have a significant positive effect on Yield2 and there was a significant negative interaction with slope in Cochno Mid (Table 5.8). The fitted regression for the Slope model is shown in Figure 5.6. As shown in the regression Table 5.9, there were no significant relationships between Aspect and the density of *F. hepatica*. The goodness of fit and the overdispersion of models was examined and documented in Table 5.10.

Table 5.7: Univariate regression model of Elevation

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-37.6058	27.3979	-1.3726	0.1763
CochnoLow	55.7872	39.6856	1.4057	0.1662
CochnoMid	27.2055	33.3628	0.8154	0.4188
DumgoyneNorth	28.7107	29.6251	0.9691	0.3373
DumgoyneSouth	48.8199	30.8917	1.5804	0.1206
Elevation	0.1136	0.0942	1.2068	0.2334
CochnoLow:Elevation	-0.2988	0.2551	-1.1714	0.2472
CochnoMid:Elevation	-0.0739	0.1477	-0.5002	0.6192
DumgoyneNorth:Elevation	0.0491	0.4682	0.1049	0.9169
DumgoyneSouth:Elevation	-0.7567	0.5933	-1.2754	0.2083

Table 5.8: Univariate regression model of Slope

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-6.0409	1.0146	-5.9538	0.000000
CochnoLow	2.2600	1.2940	1.7466	0.0871
CochnoMid	3.3780	1.3081	2.5823	0.0129
DumgoyneNorth	0.7798	1.4686	0.5310	0.5979
DumgoyneSouth	2.5229	1.3996	1.8026	0.0777
Slope	0.2436	0.1354	1.7985	0.0784
CochnoLow:Slope	-0.4891	0.3941	-1.2409	0.2207
CochnoMid:Slope	-0.6134	0.3225	-1.9018	0.0632
DumgoyneNorth:Slope	-0.0934	0.4119	-0.2267	0.8216
DumgoyneSouth:Slope	-1.7799	1.3154	-1.3531	0.1824

Table 5.9: Univariate regression model of Aspect

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-5.5524	2.7578	-2.0134	0.0507
CochnoLow	2.0106	1.4877	1.3515	0.1839
CochnoMid	2.7914	1.4115	1.9776	0.0547
DumgoyneNorth	0.0668	1.8078	0.0369	0.9707
DumgoyneSouth	1.5112	1.4848	1.0178	0.3148
NorthEast	-0.7295	2.6725	-0.2730	0.7863
NorthWest	2.7643	2.8164	0.9815	0.3321
SouthEast	0.3605	2.8405	0.1269	0.8996
SouthWest	-0.4816	2.3951	-0.2011	0.8416
West	-0.7515	3.6717	-0.2047	0.8388
CochnoMid:NorthEast	-0.1365	1.4952	-0.0913	0.9277
DumgoyneNorth:NorthEast	1.1755	2.3652	0.4970	0.6218
DumgoyneNorth:NorthWest	-3.3031	3.3166	-0.9959	0.3251
DumgoyneSouth:NorthWest	-3.5948	1.7670	-2.0344	0.0484
CochnoLow:SouthEast	-1.2084	1.6853	-0.7170	0.4774
CochnoMid:SouthEast	-1.8535	1.6248	-1.1408	0.2606
DumgoyneNorth:SouthEast	1.4990	2.0210	0.7417	0.4625

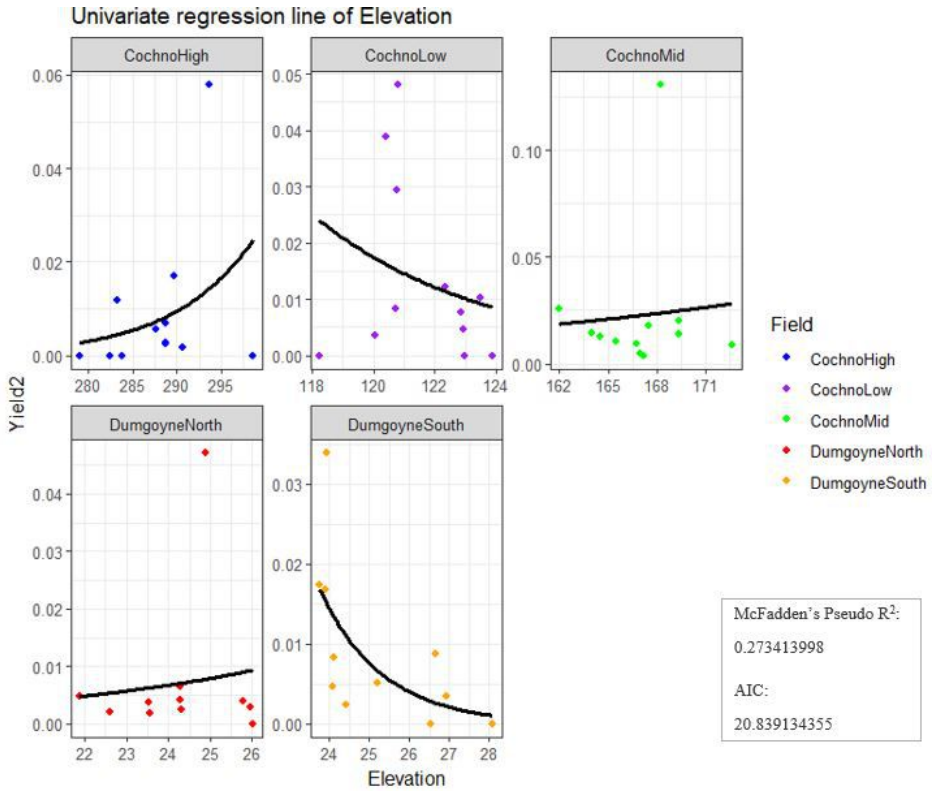


Figure 5.5: The dot plot and fitted line of Elevation against Yield2 for each field.

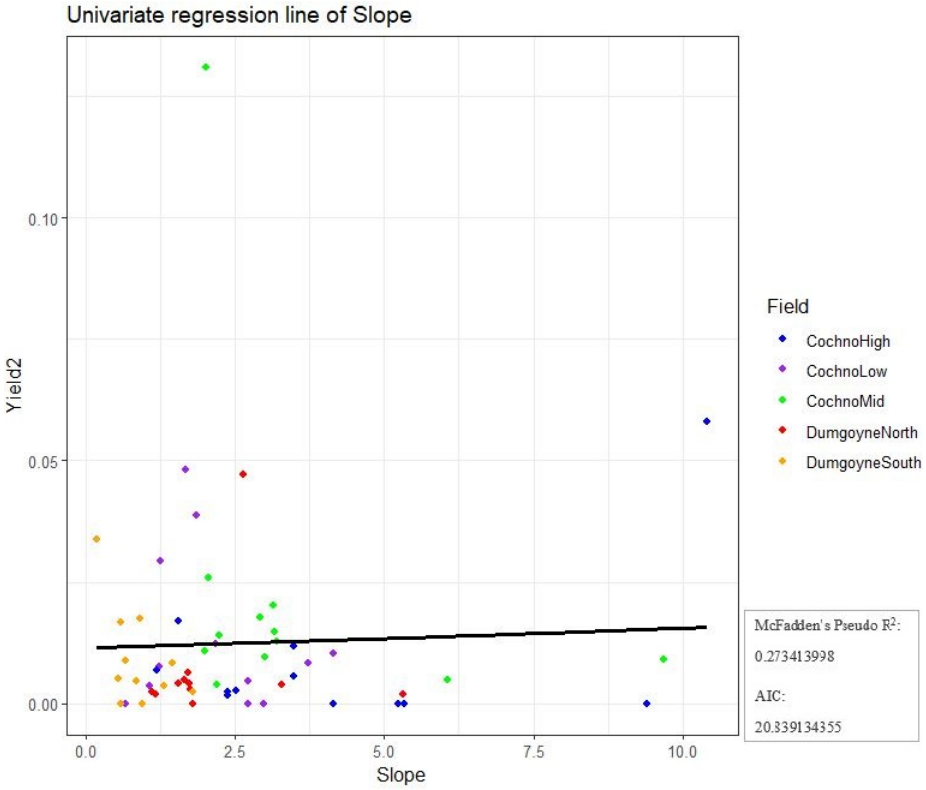


Figure 5.6: The dot plot and fitted line of Slope against Yield2 for each field.

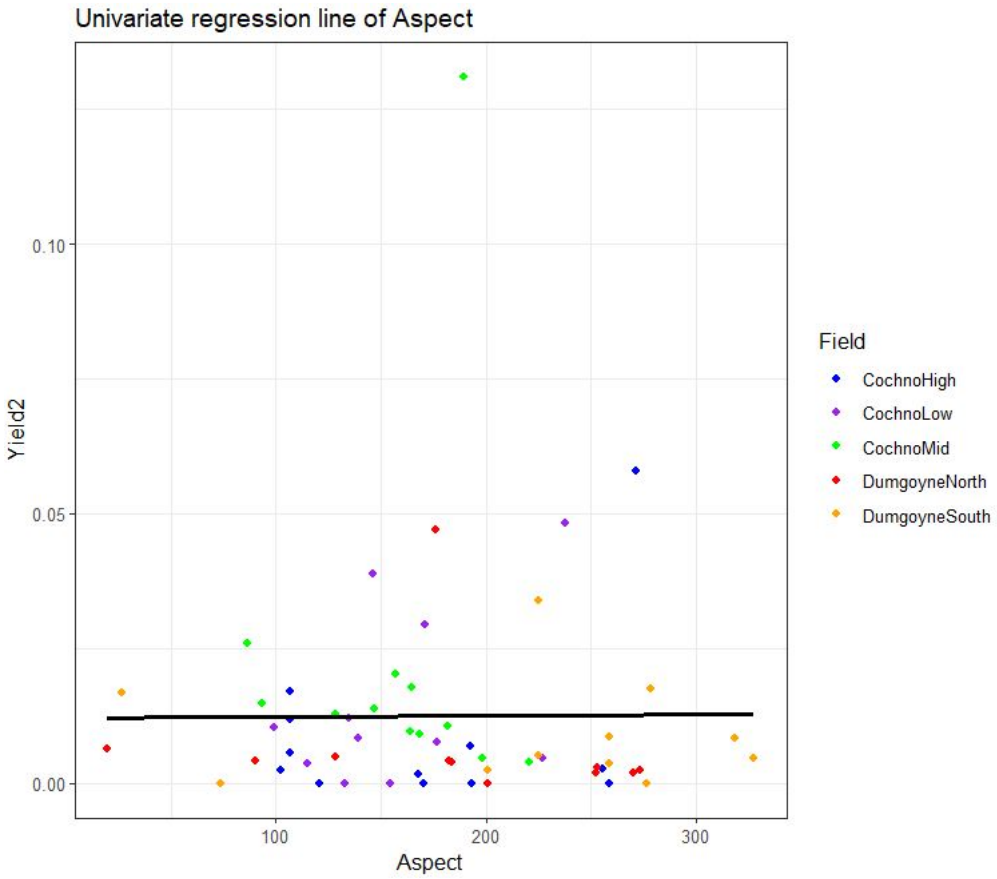


Figure 5.7: The dot plot of Aspect against Yield2.

Table 5.10: Chi-squared goodness of fit for topographical factors models

Models	Variables	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
Yield2~Field *Elevation	NULL			57	1.1549	
	Field	4	0.139	53	1.016	0.247
	Elevation	1	0.011	52	1.004	0.510
	Field:Elevation	4	0.096	48	0.908	0.443
Yield2~Field *Slope	NULL			57	1.1549	
	Field	4	0.139	53	1.016	0.164
	Slope	1	0.0001	52	1.015	0.954
	Field:Slope	4	0.176	48	0.839	0.083
Yield2~Field *Aspect	NULL			57	1.1549	
	Field	4	0.139	53	1.016	0.099
	Aspect	5	0.062	48	0.954	0.628
	Field:Aspect	7	0.256	41	0.698	0.045

5.1.4 Modeling pasture grass against *F. hepatica* density in pasture

As mentioned in Chapter 3.2, the forage data shows collinearity and therefore, the univariate regression model was built on Grass Height and Grass Weight separately. The Forage factors were nested within Field. Cochno High was used as the reference group. There were two models:

- *Yield2* $\sim$ *Field * GrassHeight*
- *Yield2* $\sim$ *Field * GrassWeight*

Including Grass Height in the first model, shows that none of the covariates were significantly associated with Yield2 (Table 5.11). On the other hand, the univariate regression model of Grass Weight demonstrated that just the Grass Weight in Dumgoyne North displayed a significantly positive relationship between Grass Weight and Yield2 (Table 5.12). Figure 5.8 and Figure 5.9 show the relationship of Grass Height and Yield2, with fitted lines of two models. The chi-squared goodness of fit for forage models is shown in Table 5.13, which shows that the Grass Weight model might fit better than Grass Height model.

Table 5.11: Univariate regression model of Grass Height

	Estimate	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept)	-12.0873	5.5662	-2.1716	0.0349
CochnoLow	9.4144	5.7775	1.6295	0.1098
CochnoMid	9.2738	5.6491	1.6417	0.1072
DumgoyneNorth	5.1934	7.2061	0.7207	0.4746
DumgoyneSouth	10.6099	6.1699	1.7196	0.0919
GrassHeight	0.3886	0.2801	1.3877	0.1716
CochnoLow:GrassHeight	-0.5317	0.3131	-1.6979	0.0960
CochnoMid:GrassHeight	-0.4912	0.2990	-1.6428	0.1070
DumgoyneNorth:GrassHeight	-0.2990	0.3461	-0.8640	0.3919
DumgoyneSouth:GrassHeight	-0.6220	0.3453	-1.8014	0.0779

Table 5.12: Univariate regression model of Grass Weight

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-3.8531	0.9757	-3.9489	0.0003
CochnoLow	1.4667	1.2737	1.1515	0.2552
CochnoMid	2.5696	1.6528	1.5546	0.1266
DumgoyneNorth	-11.1054	5.7080	-1.9456	0.0576
DumgoyneSouth	2.7059	2.7347	0.9895	0.3274
GrassWeight	-0.0023	0.0025	-0.8953	0.3751
CochnoLow:GrassWeight	-0.0059	0.0046	-1.2819	0.2060
CochnoMid:GrassWeight	-0.0061	0.0053	-1.1465	0.2573
DumgoyneNorth:GrassWeight	0.0163	0.0079	2.0801	0.0429
DumgoyneSouth:GrassWeight	-0.0088	0.0088	-0.9986	0.3230

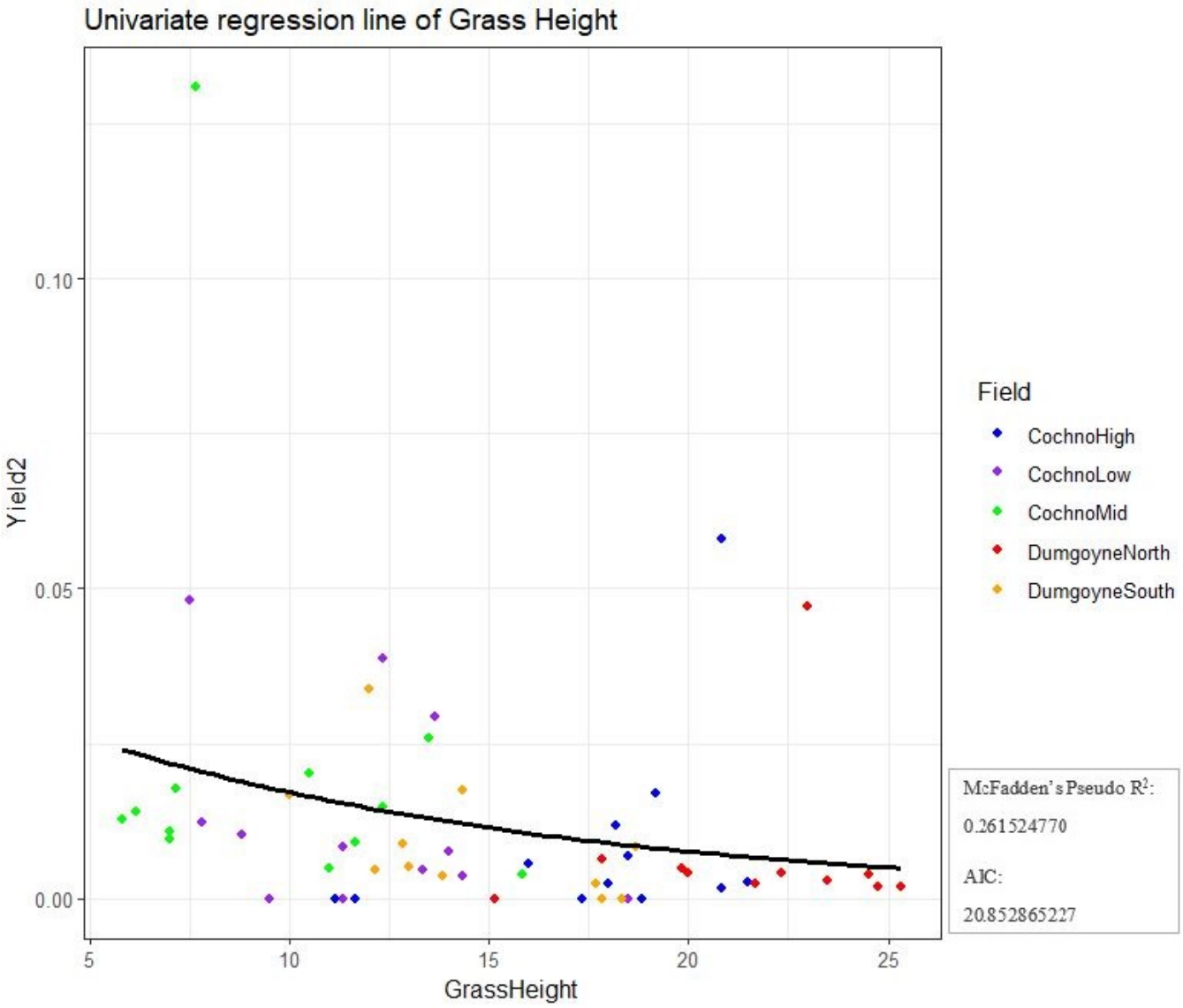


Figure 5.8: The dot plot and fitted line of Grass Height against Yield2.

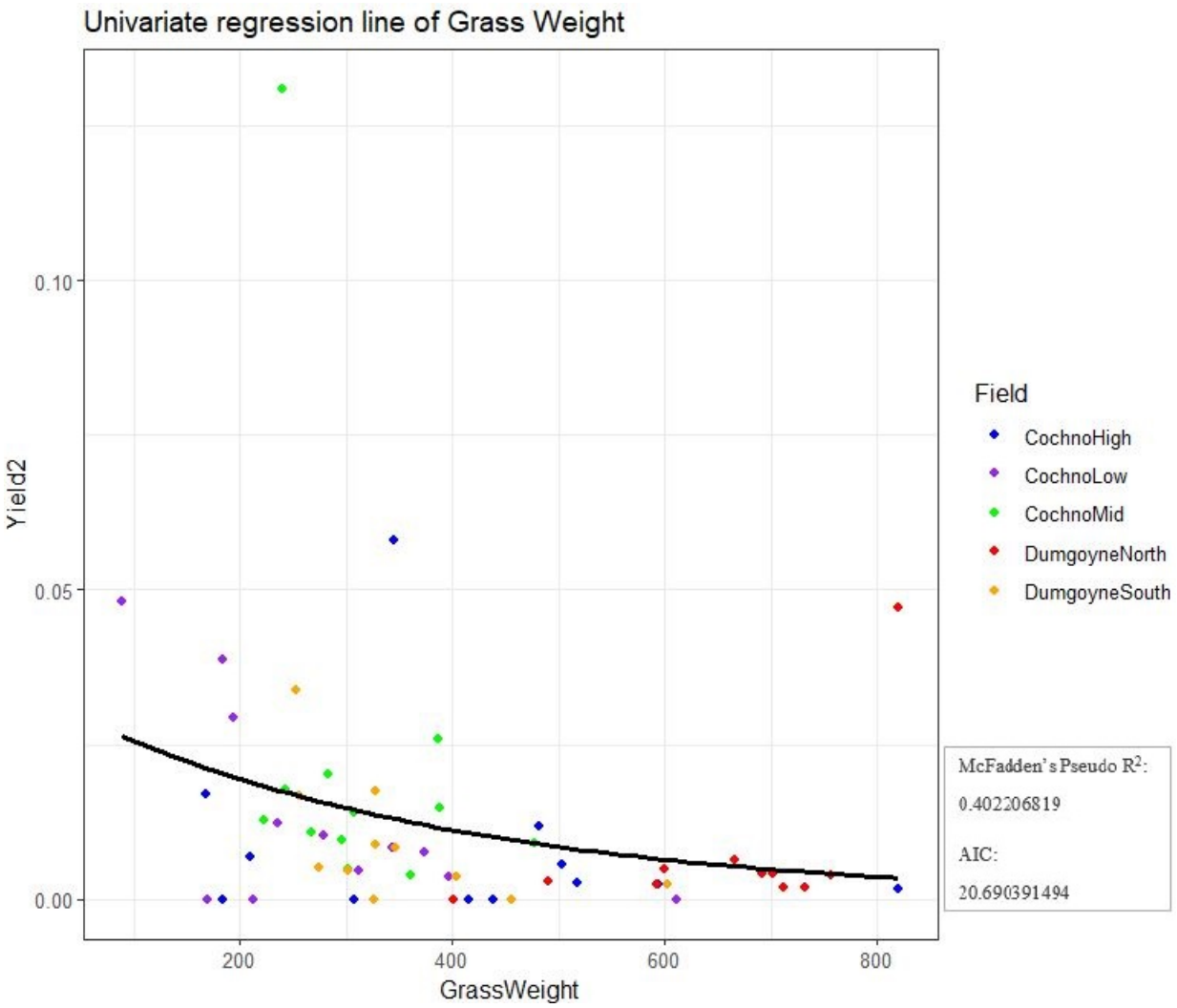


Figure 5.9: The dot plot and fitted line of Grass Weight against Yield2.

Table 5.13: Chi-squared goodness of fit for forage models

Models	Variables	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
Yield2~Field *GrassHeight	NULL			57	1.1549	
	Field	4	0.139	53	1.016	0.207
	GrassHeight	1	0.015	52	1.000	0.418
	Field:GrassHeight	4	0.147	48	0.853	0.182
Yield2~Field *GrassWeight	NULL			57	1.1549	
	Field	4	0.139	53	1.016	0.092
	GrassWeight	1	0.089	52	0.926	0.024
	Field:GrassWeight	4	0.236	48	0.690	0.009

5.2 Multivariate regression model

To deal with multicolinearity, only the more significant variables of a group of related variables were selected for a multivariate regression model. According to Section 4.1, there are four significant variables, NDVI, Grass Weight, Soil Moisture and Slope, which were included in the multivariate regression model Model 1 as the nested factors within fields. Yield2 was the response variable.

Model 1: *Yield2 ~ Field * (NDVI + GrassWeight + SoilMoisture + Slope)*

Table 5.14 lists the coefficients of regression model M1. Most covariates were not significant at the 95% confidence interval. However, there were some factors which show a significant relationship with the density of *F. hepatica* in pasture. Cochno Mid had a significantly higher density and Soil Moisture had a significant positive association with Yield2 overall, but had a negative association with Yield2 in the Cochno Mid and Dumgoyne North. NDVI values in Cochno Mid also had a significant negative relationship with the Yield2. Table 5.15 shows the chi-squared goodness of fit test of Model 1, which showed that most covariates fitted the model well, apart from Slope.

Table 5.14: Multivariate regression model

	Estimate	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept)	-17.066	6.168	-2.767	0.009
CochnoLow	10.249	7.758	1.321	0.196
CochnoMid	25.025	6.526	3.835	0.001
DumgoyneNorth	56.912	70.044	0.813	0.422
DumgoyneSouth	3.596	16.719	0.215	0.831
NDVI	7.764	7.023	1.106	0.277
GrassWeight	-0.0002	0.003	-0.077	0.939
SoilMoisture	19.222	7.812	2.461	0.019
Slope	-0.320	0.222	-1.442	0.159
CochnoLow:NDVI	-11.112	7.313	-1.520	0.138
CochnoMid:NDVI	-22.260	8.368	-2.660	0.012
DumgoyneNorth:NDVI	-62.385	81.747	-0.763	0.451
DumgoyneSouth:NDVI	-4.263	17.702	-0.241	0.811
CochnoLow:GrassWeight	-0.005	0.004	-1.179	0.247
CochnoMid:GrassWeight	-0.003	0.005	-0.559	0.580
DumgoyneNorth:GrassWeight	0.012	0.007	1.680	0.102
DumgoyneSouth:GrassWeight	-0.002	0.009	-0.194	0.848
CochnoLow:SoilMoisture	-11.118	9.769	-1.138	0.263
CochnoMid:SoilMoisture	-22.179	8.845	-2.507	0.017
DumgoyneNorth:SoilMoisture	-33.525	13.336	-2.514	0.017
DumgoyneSouth:SoilMoisture	-7.235	10.810	-0.669	0.508
CochnoLow:Slope	0.749	0.435	1.721	0.095
CochnoMid:Slope	-0.023	0.301	-0.077	0.939
DumgoyneNorth:Slope	0.278	0.540	0.514	0.611
DumgoyneSouth:Slope	-0.804	1.128	-0.713	0.481

Table 5.15: Chi-squared goodness of fit for multivariate regression model

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			57	1.155	
Field	4	0.139	53	1.016	0.002
NDVI	1	0.089	52	0.927	0.001
GrassWeight	1	0.065	51	0.862	0.004
SoilMoisture	1	0.033	50	0.828	0.042
Slope	1	0.001	49	0.827	0.679
Field:NDVI	4	0.298	45	0.528	0.00000
Field:GrassWeight	4	0.149	41	0.379	0.001
Field:SoilMoisture	4	0.095	37	0.284	0.018
Field:Slope	4	0.038	33	0.246	0.314

Chapter 6

Discussion and conclusion

6.1 Discussion

The main objective of this study was to apply earth observation and GIS approaches to detect the high risk areas for *F. hepatica* within a farm and try to predict the geographical distribution of *F. hepatica*. This study also explored the relationships among environment and the density *F. hepatica* in pasture at the field-level within farms.

Firstly, *F. hepatica* metacercariae were found in each field and some locations had a high number of metacercarial cysts. Yield, which was calculated as the number of metacercarial cysts per gram of pasture grass, was used to quantify the *F. hepatica* density in forage. Higher yield values are likely to result in higher risk of *F. hepatica* infection in sheep. The results showed that the density of metacercariae in pasture was significantly different between the five fields, but the pattern of distribution of counts in the five fields was similar. Most of the grass samples were clean or had low counts of metacercariae, but a small number of samples had high metacercarial cyst counts of *F. hepatica*.

The environmental factors including elevation, slope and soil were also compared among five fields and showed significant differences among fields. Soil moisture and soil temperature differed significantly among fields. The soil moisture of most sampling locations was high, most exceeding 50%, which is an ideal environmental condition for the life-cycle of *F. hepatica*. The average soil temperature was 15°C, which is within the ideal temperature range for cercarial encystment. In addition, to conventionally measured environmental factors, the normalized difference vegetation index (NDVI) and normalized difference water index (NDWI) generated from Sentinel-2 satellite images also differed significantly among the five fields. The NDVI values were all high, most values exceeding 0.5, meaning that there was abundant pasture grass. In contrast, the NDWI values were not high, all values being less than 0.5 with some values below zero, indicating very dry areas or vegetation with very low water content which gave different

information from the *in situ* soil moisture measurements. Grass height and grass weight were used as factors to indicate amounts of forage.

The analysis of the relationships among environmental factors was conducted at the farm-level. The NDWI and NDVI showed significant positive relationships with each other within all five fields. The abundance of pasture showed significant associations with the terrain factors, such as elevation and slope, in all five fields. For example, in Dumgoyne South, both grass height and grass weight were positively related to the slope, and in Dumgoyne North, the grass height and aspect were also positively associated. Also, at Cochno, the grass height of Cochno High was positively related with the elevation, while the grass weight of Cochno Mid was positively associated with the slope. Besides, there were significant connections between soil properties and topographical factors in the five fields. First, the soil moisture had a negative relationship with elevation in Dumgoyne South and Cochno Low. The Cochno Mid also showed the negative relationships between soil moisture and slope. The only field where soil moisture had a positive relationship to the terrain factors and slope was Cochno High. As for the soil temperature in each field, there were negative relationships with aspect and elevation in the Cochno High, and there was also a positive relationship with altitude in Cochno Mid. The reason for this might be the outh facing areas will receive more sunlight and could have higher grass heights.

Application of remote sensing methods to detect land surface was one of the objectives in this study. The Sentinel-2 images of 31 August 2017 were used in this study. In Cochno High, the NDVI showed a strong positive association with grass height. In Cochno Mid, the NDWI was negatively associated with grass weight. Furthermore, in Dumgoyne South, both NDWI and NDVI were positively related to soil temperature. It is worth noting that negative associations between soil moisture and both NDWI and NDVI were also identified. A similar situation emerged in Cochno Low, where soil temperature was negatively associated with both NDWI and NDVI. According to the related research, the NDWI were used to detect the water bodies from vegetation or other background land surface, and also can indicate the water content of vegetation (Soti et al., 2009; Xu, 2006; De Roeck et al., 2014). The NDVI were normally used to indicate the abundance of pasture vegetation (Gao, 1996; Manyangadze et al., 2016; De Roeck et al., 2014). However, the results of this study show that the remote sensing indices did not reflect the actual land situation. This contradiction between the retrieved values from satellite images and the measured values from the field survey was found in all five fields. Possible reasons for this contradiction could be the choice of spectral bands and the spectral resolution. The NDWI in this study used the band combination of NIR (690 nm to 980 nm) and SWIR (1470 nm to 1750 nm), which can distinguish water with soil from other type of land cover types, but cannot separate water from vegetation (Soti et al., 2009).

The univariate regression model of each environmental variable was established against the Yield2, the number of metacercarial cysts per 0.1 gram, to explain the relationships among environment variables and the density of *F. hepatica* in pasture. The regression results showed that few univariate models had significant coefficients of covariates and there were only a few variables in specific fields having a significant association with the Yield2. This means the relationships among environment variables and *F. hepatica* were weak; and the directions of the relation- ships differed among fields. This results suggest that remote sensing and environmental factors ,such as terrain factors and soil texture, might not be reliable when monitoring *F. hepatica* in farm-level, especially in the farm with high transmission of fasciolosis. This lack of association could be explained by the following factors. The first factor could be the Validity of the environmental observations. The measurements taken in the field are single time point observations and this does not deal well with the phenology of pasture in relation to collection of Sentinel 2 data. The field survey were conducted on 5 and 6 September 2017, while the NDWI and NDVI data were obtained from the satellite images of 31 August 2017. Temporal changes might affect the measurements of grass, soil moisture and temperature, causing remote sensing indices are different with real measurements. In this study, the date of satellite images was six days ahead of the field survey and the soil texture and vegetation are very likely to change in these days. This might cause what the remote sensing indices shown is different with the field survey data. Moreover, variation in recovery rates of metacercariae can also impact the results. In this study, seven replicates were carried out for each sample. At the beginning of the lab analysis, some samples had 10 replicates and showed a difference compared with those with seven replicates. This might lead to the loss of metacercarial cysts. Sheep density and farm size might cause the different directions of relationships among fields. In addition, the remote sensing indices used in this study cannot reflect *in situ* measurements. The spatial resolutions of Sentinel-2 images are 10 m and 20 m, and in this study, the 10 metre band was resampled to 20 metres resolution for the NDWI value calculations, which means the area of a single pixel is 400 m^2. The variation of *F. hepatica* within this area can be ignored. Moreover, the remote sensing indices may not detect moisture areas.

6.2 Conclusion

In conclusion, this study mapped the environment data to the *F. hepatica* risk and assessed the variation of environmental factors among fields. The data used in this study were from multiple sources, including satellite images, LiDAR images, field survey measurements. Four categories including nine variables of environmental factors, remote sensing indices, topographical factors, soil properties, and grazing grass, were used in the model. This study explored the impact of environmental factors on the geographical distribution of *F. hepatica*.

The results of the regression model demonstrated that some environmental factors were significantly related to the *F. hepatica* in particular areas. The negative binomial regression model employed by this study performed well, and the multicollinearity of the multivariate model was reduced by selecting the key parameters. In the NDVI model, intercept and Cochno Mid had significant coefficients. The estimate of Cochno Mid was 12.41 which meant it had a positive relationship with Yield2. Soil moisture model had the most significant factors. When modelling soil moisture to Yield2, field Cochno Mid and Dumgoyne North had significant, positive regression coefficients. The Soil moisture itself also showed a significant relationship with the Yield2, which was also a positive relation. Another univariate model with significant factors is the Slope model, in which the Cochno Mid and the intercept were significant factors. The dependent variables used in the multivariate regression model were NDVI, Soil Moisture, Slope, and Grass Weight. The result of regression demonstrated that there were restricted relations between environmental factors and *F. hepatica*. The Cochno Mid was the only field which was significantly related to the Yield2, with a positive estimate coefficient value implying the positive association with the risk of infection of *F. hepatica*. Soil moisture was the only environmental variable which had a significant coefficient in the multivariate model. The coefficient of soil moisture was 19.222, which meant there was a positive relationship between soil moisture and the Yield2 across all five fields.

However, this study still has limitations. First, the use of satellite images needs to be developed. Farm with small area (under 20ha), mid resolution images is too rough and too wide to show detailed land surface information. Higher spatial resolution images can provide more detailed land surface information because each farm area can be covered by more images. It is worth using different remote sensing indices, such as modified normalized water index and normalized difference turbidity index, which can detect and distinguish free standing water bodies and water turbidity (Lacaux et al., 2007; Soti et al., 2010). Future studies on mapping environment to *F. hepatica* at farm-level could try multiple sensor images. For example, the hyperspectral images such as MODIS and Earth Observing-1 can improve land cover classification and detect moisture areas. In the analysis of *F. hepatica*, it require us to distinguish moisture areas with vegetation from the vegetation alone, which causes the single cell of images contains information of various features (soil, water, vegetation, etc.) and therefore reduce the accuracy of detection. Hyperspectral images can reduce this problem (Cheng et al., 2006; Datt et al., 2003).On the other hand, repeated field measurements are also necessary. Future studies should take the dynamic nature of pasture grass into consideration.

Bibliography

Centre for environmental data analysis. http://www.ceda.ac.uk/.

Digimap lidar. http://digimap.edina.ac.uk/lidar, a.

Digimap aerial. http://digimap.edina.ac.uk/aerial, b.

Earth engine code editor. https://code.earthengine.google.com/.

Hydraprobe manual. https://www.stevenswater.com/products/hydraprobe/.

Nadis parasite forecast - october 2017. http://webinars.nadis.org.uk/media/42985/17-10_parasite_forecast__sqp_.pdf.

Mouna Abrous, Daniel Rondelaud, Gilles Dreyfuss, and Jacques Cabaret. Infection of Lymnaea truncatula and Lymnaea glabra by Fasciola hepatica and Paramphistomum daubneyi in farms of central France. *Veterinary Research*, 30(1):113–118, 1999.

European Space Agency. Sentinel-2 msi technical guides. https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi.

Hirotugu Akaike. A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974. ISSN 15582523. doi: 10.1109/TAC.1974.1100705.

Hirotugu Akaike. Factor analysis and AIC. In *Selected Papers of Hirotugu Akaike*, pages 371–386. Springer, 1987.

D E Alsdorf, E Rodriguez, and Dennis P Lettenmaier. Measuring surface water from space. *Reviews of Geophysics*, 45(2):1–24, 2007. ISSN 87551209. doi: 10.1029/2006RG000197.1.INTRODUCTION. URL http://www.agu.org/pubs/crossref/2007/2006RG000197.shtml.

Keyhan Ashrafi, M. Adela Valero, Raquel V. Peixoto, Patricio Artigas, Miroslava Panova, and Santiago Mas-Coma. Distribution of Fasciola hepatica and F. gigantica in the endemic area of Guilan, Iran: Relationships between zonal overlap and phenotypic traits. *Infection, Genetics and Evolution*, 31:95–109, apr 2015. ISSN 1567-1348. doi: 10.1016/J.MEEGID.2015.

01.009. URL https://www.sciencedirect.com/science/article/pii/S1567134815000131.

Brian William Avery et al. Soil classification for england and wales [higher categories]. *Soil classification for England and Wales [higher categories].*, (14), 1980.

Maurice Stevenson Bartlett. Properties of sufficiency and statistical tests. *Proc. R. Soc. Lond. A*, 160(901):268–282, 1937. ISSN 0080-4630.

N. J. Beesley, C. Caminade, J. Charlier, R. J. Flynn, J. E. Hodgkinson, A. Martinez-Moreno, M. Martinez-Valladeres, J. Perez, L. Rinaldi, and D. J. L. Williams. Fasciola and fascioliosis in ruminants in Europe: Identifying research needs. *Transboundary and Emerging Diseases*, (April):1–18, 2017. ISSN 18651674. doi: 10.1111/tbed.12682. URL http://doi.wiley.com/10.1111/tbed.12682.

S. Bennema, J. Vercruysse, E. Claerebout, T. Schnieder, C. Strube, E. Ducheyne, G. Hendrickx, and J. Charlier. The use of bulk-tank milk ELISAs to assess the spatial distribution of Fasciola hepatica, Ostertagia ostertagi and Dictyocaulus viviparus in dairy cattle in Flanders (Belgium). *Veterinary Parasitology*, 165(1-2):51–57, 2009. ISSN 03044017. doi: 10.1016/j.vetpar.2009.07.006.

S. C. Bennema, E. Ducheyne, J. Vercruysse, E. Claerebout, G. Hendrickx, and J. Charlier. Relative importance of management, meteorological and environmental factors in the spatial distribution of Fasciola hepatica in dairy cattle in a temperate climate zone. *International Journal for Parasitology*, 41(2):225–233, 2011. ISSN 00207519. doi: 10.1016/j.ijpara.2010.09.003. URL http://dx.doi.org/10.1016/j.ijpara.2010.09.003.

Joseph C Boray. Experimental fascioliasis in Australia. In *Advances in parasitology*, volume 7, pages 95–210. Elsevier, 1969. ISBN 0065-308X.

Antonio Bosco, Laura Rinaldi, Vincenzo Musella, Alessandra Amadesi, and Giuseppe Cringoli. Outbreak of acute fascioliosis in sheep farms in a Mediterranean area arising as a possible consequence of climate change. *Geospatial Health*, 9(2), 2015. ISSN 19707096. doi: 10.4081/gh.2015.354.

Pietro Ceccato, Stéphane Flasse, Stefano Tarantola, Stéphane Jacquemoud, and Jean-Marie Grégoire. Detecting vegetation leaf water content using reflectance in the optical domain. *Remote sensing of environment*, 77(1):22–33, 2001. ISSN 0034-4257.

Johannes Charlier, Karen Soenen, Els De Roeck, Wouter Hantson, Els Ducheyne, Frieke Van Coillie, Robert De Wulf, Guy Hendrickx, and Jozef Vercruysse. Longitudinal study on the temporal and micro-spatial distribution of Galba truncatula in four

farms in Belgium as a base for small-scale risk mapping of Fasciola hepatica. *Parasites & vectors*, 7:528, 2014a. ISSN 1756-3305. doi: 10.1186/s13071-014-0528-0. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4247775{&}tool=pmcentrez{&}rendertype=abstract.

Johannes Charlier, Jozef Vercruysse, E Morgan, J Van Dijk, and D J L Williams. Recent advances in the diagnosis, impact on production and prediction of Fasciola hepatica in cattle. *Parasitology*, 141(3):326–335, 2014b. ISSN 0031-1820.

Johannes Charlier, Aklilu H. Ghebretinsae, Bruno Levecke, Els Ducheyne, Edwin Claerebout, and Jozef Vercruysse. Climate-driven longitudinal trends in pasture-borne helminth infections of dairy cattle. *International Journal for Parasitology*, 46(13-14):881–888, dec 2016. ISSN 0020-7519. doi: 10.1016/J.IJPARA.2016.09.001. URL https://www.sciencedirect.com/science/article/pii/S0020751916302284.

Yen-Ben Cheng, Pablo J Zarco-Tejada, David Riaño, Carlos A Rueda, and Susan L Ustin. Estimating vegetation water content with hyperspectral data for different canopy scenarios: Relationships between aviris and modis indexes. *Remote Sensing of Environment*, 105(4): 354–366, 2006.

Peter Dambach, Ali Sié, Jean Pierre Lacaux, Cécile Vignolles, Vanessa Machault, and Rainer Sauerborn. Using high spatial resolution remote sensing for risk mapping of malaria occurrence in the Nouna district, Burkina Faso. *Global Health Action*, 2(1), 2009. ISSN 16549880. doi: 10.3402/gha.v2i0.2094.

Bisun Datt, Tim R McVicar, Tom G Van Niel, David LB Jupp, and Jay S Pearlman. Preprocessing eo-1 hyperion hyperspectral data to support the application of agricultural indexes. *IEEE Transactions on Geoscience and Remote Sensing*, 41(6):1246–1259, 2003.

Els De Roeck, Frieke Van Coillie, Robert De Wulf, Karen Soenen, Johannes Charlier, Jozef Vercruysse, Wouter Hantson, Els Ducheyne, and Guy Hendrickx. Fine-scale mapping of vector habitats using very high resolution satellite imagery: A liver fluke case-study. *Geospatial Health*, 8(3):S671–S683, 2014. ISSN 19707096. doi: 10.4081/gh.2014.296.

R V Facey and P D Marsden. Fascioliasis in man: an outbreak in Hampshire. *British medical journal*, 2(5199):619, 1960.

Naomi J. Fox, Piran C.L. White, Colin J. McClean, Glenn Marion, Andy Evans, and Michael R. Hutchings. Predicting impacts of climate change on fasciola hepatica risk. *PLoS ONE*, 6(1): 19–21, 2011. ISSN 19326203. doi: 10.1371/journal.pone.0016126.

Bo Cai Gao. NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3):257–266, 1996. ISSN 00344257. doi: 10.1016/S0034-4257(96)00067-3.

Fisheries Great Britain. Ministry of Agriculture and Food. *Manual of Veterinary Parasitological Laboratory Techniques*. Reference Book - Ministry of Agriculture, Fisheries and Food. H.M. Stationery Office, 1986. ISBN 9780112427247. URL https://books.google.co.uk/books?id=sLs4AQAAIAAJ.

W Gregory and D Foreman. Nonparametric Statistics for Non-Statisticians. *Hoboken: John Wiley and Sons*, 2009.

Christopher Hackney and Alexander Clayton. Unmanned aerial vehicles (uavs) and their application in geomorphic mapping. 2015.

Meghan Halabisky. Object-based classification ofâĂL'semi-arid wetlands. *Journal of Applied Remote Sensing*, 5(1):053511, 2011. ISSN 1931-3195. doi: 10.1117/1.3563569. URL http://remotesensing.spiedigitallibrary.org/article.aspx?doi=10.1117/1.3563569.

John Michael Hodgson. Soil survey field handbook describing and sampling soil profiles. *Technical monograph*, 1974.

Alison Howell, Matthew Baylis, Rob Smith, Gina Pinchbeck, and Diana Williams. Epidemiology and impact of Fasciola hepatica exposure in high-yielding dairy herds. *Preventive veterinary medicine*, 121(1-2):41–48, 2015. ISSN 0167-5877.

Abdel-Nasser A Hussein and Refaat MA Khalifa. Fascioliasis prevalences among animals and human in upper egypt. *Journal of King Saud University-Science*, 22(1):15–19, 2010.

Ishmael Festus Jaja, Borden Mushonga, Ezekiel Green, and Voster Muchenje. Seasonal prevalence, body condition score and risk factors of bovine fasciolosis in south africa. *Veterinary and Animal Science*, 4:1–7, 2017.

Vaia Kantzoura, Marc K. Kouam, Haralambos Feidas, Denitsa Teofanova, and Georgios Theodoropoulos. Geographic distribution modelling for ruminant liver flukes (Fasciola hepatica) in south-eastern Europe. *International Journal for Parasitology*, 41(7):747–753, 2011. ISSN 00207519. doi: 10.1016/j.ijpara.2011.01.006. URL http://dx.doi.org/10.1016/j.ijpara.2011.01.006.

F Kenyon, N D Sargison, P J Skuce, and F Jackson. Sheep helminth parasitic disease in south eastern Scotland arising as a possible consequence of climate change. *Veterinary parasitology*, 163(4):293–297, 2009. ISSN 0304-4017.

Muhammad Kasib Khan, Muhammad Sohail Sajid, Hasan Riaz, Nazia Ehsan Ahmad, Lan He, Muhammad Shahzad, Altaf Hussain, Muhammad Nisar Khan, Zafar Iqbal, and Junlong Zhao. The global burden of fasciolosis in domestic animals with an outlook on the contribution of new approaches for diagnosis and control. *Parasitology research*, 112(7):2421–2430, 2013.

William H Kruskal and W Allen Wallis. Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 47(260):583–621, 1952. ISSN 0162-1459.

Birte Kuerpick, Franz J Conraths, Christoph Staubach, Andreas Fröhlich, Thomas Schnieder, and Christina Strube. Seroprevalence and GIS-supported risk factor analysis of Fasciola hepatica infections in dairy herds in Germany. *Parasitology*, 140(8):1051–60, 2013. ISSN 1469-8161. doi: 10.1017/S0031182013000395. URL http://journals.cambridge.org/abstract{_}S0031182013000395.

J. P. Lacaux, Y. M. Tourre, C. Vignolles, J. A. Ndione, and M. Lafaye. Classification of ponds from high-spatial resolution remote sensing: Application to Rift Valley Fever epidemics in Senegal. *Remote Sensing of Environment*, 106(1):66–74, 2007. ISSN 00344257. doi: 10.1016/j.rse.2006.07.012.

Jonathan D LaPook, Arthur M Magun, Katherine G Nickerson, and Jay I Meltzer. Sheep, watercress, and the Internet. *The Lancet*, 356(9225):218, 2000. ISSN 0140-6736.

Damanbha Lyngdoh, Sunil Sharma, Bishnupada Roy, and Veena Tandon. Animal Fascioliasis: Perspectives from high altitudinal regions. *Veterinary Parasitology*, 232:21–31, dec 2016. ISSN 0304-4017. doi: 10.1016/J.VETPAR.2016.11.007. URL https://www.sciencedirect.com/science/article/pii/S0304401716304393.

Tawanda Manyangadze, Moses John Chimbari, Michael Gebreslasie, Pietro Ceccato, and Samson Mukaratirwa. Modelling the spatial and seasonal distribution of suitable habitats of schistosomiasis intermediate host snails using Maxent in Ndumo area, KwaZulu-Natal Province, South Africa. *Parasites & vectors*, 9(1):572, 2016. ISSN 1756-3305.

C. M. McCann, M. Baylis, and D. J L Williams. Seroprevalence and spatial distribution of Fasciola hepatica-infected dairy herds in England and Wales. *Veterinary Record*, 166(20): 612–617, 2010a. ISSN 00424900. doi: 10.1136/vr.b4836.

Catherine M. McCann, Matthew Baylis, and Diana J L Williams. The development of linear regression models using environmental variables to explain the spatial distribution of Fasciola hepatica infection in dairy herds in England and Wales. *International Journal for Parasitology*, 40(9):1021–1028, 2010b. ISSN 00207519. doi: 10.1016/j.ijpara.2010.02.009. URL http://dx.doi.org/10.1016/j.ijpara.2010.02.009.

Peter McCullagh and John A Nelder. *Generalized linear models*, volume 37. CRC press, 1989. ISBN 0412317605.

Daniel McFadden. Conditional logit analysis of qualitative choice behavior. 1973.

Khalid Mehmood, Hui Zhang, Ahmad Jawad Sabir, Rao Zahid Abbas, Muhammad Ijaz, Aneela Zameer Durrani, Muhammad Hassan Saleem, Mujeeb Ur Rehman, Muhammad Kashif Iqbal, Yajing Wang, Hafiz Ishfaq Ahmad, Tariq Abbas, Riaz Hussain, Muhammad Taslim Ghori, Sadaqat Ali, Aman Ullah Khan, and Jiakui Li. A review on epidemiology, global prevalence and economical losses of fasciolosis in ruminants. *Microbial Pathogenesis*, 109:253–262, aug 2017. ISSN 0882-4010. doi: 10.1016/J.MICPATH.2017.06.006. URL https://www.sciencedirect.com/science/article/pii/S0882401017304795.

Met-office. 2017 weather summaries. https://www.metoffice.gov.uk/climate/uk/summaries/2017, a.

Met-office. Climate summaries - met office. https://www.metoffice.gov.uk/climate/uk/summaries, b.

George Mitchell. Update on fasciolosis in cattle and sheep. *In Practice*, 24(7):378, 2002. ISSN 0263-841X.

NADIS. http://www.nadis.org.uk/.

NADIS. Nadis parasite forecast 2017 august. http://webinars.nadis.org.uk/media/42533/17-08_parasite_forecast__sqp_.pdf, 2017.

Adam Novobilský, Annie Engström, Sofia Sollenberg, Katarina Gustafsson, David A. Morrison, and Johan Höglund. Transmission patterns of Fasciola hepatica to ruminants in sweden. *Veterinary Parasitology*, 203(3-4):276–286, 2014. ISSN 18732550. doi: 10.1016/j.vetpar.2014.04.015.

Abbey Olsen, Klaas Frankena, Rene' Bødker, Nils Toft, Stig M Thamsborg, Heidi L Enemark, and Tariq Halasa. Prevalence, risk factors and spatial analysis of liver fluke infections in Danish cattle herds. *Parasites & vectors*, 8(1):160, 2015. ISSN 1756-3305. doi: 10.1186/s13071-015-0773-x. URL http://www.scopus.com/inward/record.url?eid=2-s2.0-84925649670{&}partnerID=tZOtx3y1.

Stacy L Ozesmi and Marvin E Bauer. Satellite remote sensing of wetlands. *Wetlands ecology and management*, 10(5):381–402, 2002. ISSN 0923-4861.

Karl Pearson. On the Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables is Such that it Can be Reasonably Supposed to have Arisen from Random Sampling. In *Breakthroughs in Statistics*, pages 11–28. Springer, 1992.

Nathalie Pettorelli. *The normalized difference vegetation index*. Oxford University Press, Oxford, first edition, 2013. ISBN 0191810142;9780191810145;.

Sorin C Popescu. Estimating biomass of individual pine trees using airborne lidar. *Biomass and Bioenergy*, 31(9):646–655, 2007. ISSN 0961-9534.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018. URL `https://www.R-project.org/`.

N. D. Sargison and P. R. Scott. Diagnosis and economic consequences of triclabendazole resistance in Fasciola hepatica in a sheep flock in south-east Scotland. *Veterinary Record*, 168(6): 159, 2011. ISSN 00424900. doi: 10.1136/vr.c5332.

University of Glasgow School of Veterinary Medicine. Cochno farm and research centre. `https://www.gla.ac.uk/schools/vet/scottishcentreforproductionanimalhealthfoodsafety/cochnofarmandresearchcentre/cochnofarm/`.

Samuel Sanford Shapiro and Martin B Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611, 1965. ISSN 0006-3444.

Philip Skuce, Jan van Dijk, Daniel Smith, and Eric Morgan. Implications of extreme weather events for risk of fluke infection. *The Veterinary record*, 175(8):198–200, aug 2014. ISSN 2042-7670. doi: 10.1136/vr.g5377. URL `http://www.ncbi.nlm.nih.gov/pubmed/25172650`.

V. Soti, C. Puech, D. Lo Seen, A. Bertran, C. Vignolles, B. Mondet, N. Dessay, and A. Tran. The potential for remote sensing and hydrologic modelling to assess the spatio-temporal dynamics of ponds in the Ferlo Region (Senegal). *Hydrology and Earth System Sciences*, 14(8):1449–1464, 2010. ISSN 10275606. doi: 10.5194/hess-14-1449-2010.

Valérie Soti, Annelise Tran, Jean Stéphane Bailly, Christian Puech, Danny Lo Seen, and Agnes Bégué. Assessing optical earth observation systems for mapping and monitoring temporary ponds in arid areas. *International Journal of Applied Earth Observation and Geoinformation*, 11(5):344–351, 2009. ISSN 15698432. doi: 10.1016/j.jag.2009.05.005.

M. A. Taylor, R. L. Coop, and Richard Wall. *Veterinary parasitology*. John Wiley and Sons, Inc, Ames, Iowa;Chichester, West Sussex;, fourth edition, 2016. ISBN 0470671629;9780470671627;.

PAUL Torgerson and JOHN Claxton. Epidemiology and control. *Fasciolosis*, 113:149, 1999.

S Tum, ML Puotinen, and DB Copeman. A geographic information systems model for mapping risk of fasciolosis in cattle and buffaloes in cambodia. *Veterinary Parasitology*, 122(2):141–149, 2004.

Cécile Vignolles, Yves M. Tourre, Oscar Mora, Laurent Imanache, and Murielle Lafaye. TerraSAR-X high-resolution radar remote sensing: An operational warning system for Rift Valley fever risk. *Geospatial Health*, 5(1):23–31, 2010. ISSN 18271987. doi: 10.4081/gh.2010.184.

WHO. Progress in assessment of morbidity due to fasciola hepatica infection : a review of recent literature. `http://apps.who.int/iris/bitstream/handle/10665/61432/WHO_SCHISTO_90.104.pdf?sequence=1&isAllowed=y`, 1990.

Hanqiu Xu. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27(14):3025–3033, 2006. ISSN 13665901. doi: 10.1080/01431160600589179.

Appendix A

Monthly precipitation and temperature of the study areas

Recent research proved demonstrated that climate data such as temperature and precipitation play key roles in the life cycle of *F. hepatica* especially in the embryonation stage and also cercariae. The climate data can be acquired from the Centre for Environmental Data Analysis (CEDA), which makes publically available the hourly rainfall and daily temperature onto its own online platform (http://www.ceda.ac.uk/). The unit of hourly rainfall is mm and the unit of daily temerature is Celsius Degree. The climate data using in this study were collected from weather station Mugdock Park (id is 954) and covers the period of from 1 January 2016 to 30 October 2017.

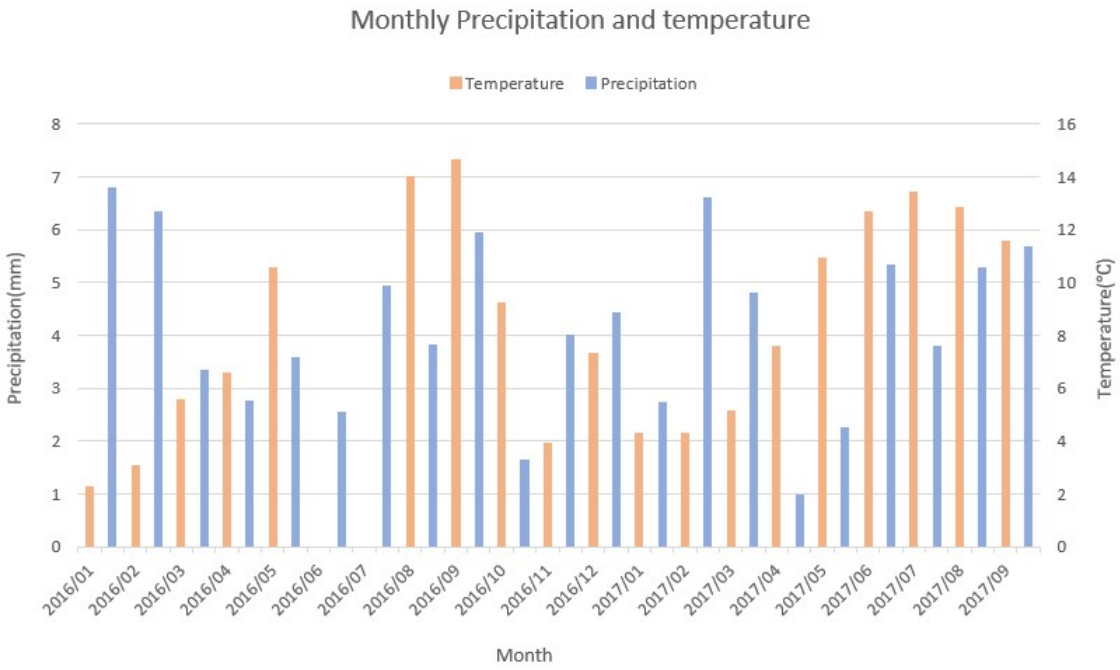


Figure A.1: Monthly precipitation (mm) and temperature (degrees Celsius) of the study areas

Appendix B

Sampling Locations of each field

The latitude and longitude of each of the points were recorded in a hand-held GPS device under the UTM/WGS84.

Table B.1: Sampling locations of Cochno

Field	Samples	Lat	Lon	Field	Samples	Lat	Lon
CochnoLow	L01	55.934637	-4.404783	CochnoMid	M01	55.9406	-4.40617
	L02	55.934043	-4.406531		M02	55.94039	-4.40617
	L03	55.934535	-4.404249		M03	55.94018	-4.406
	L04	55.934041	-4.406346		M04	55.93994	-4.4058
	L05	55.933979	-4.406023		M05	55.93964	-4.40556
	L06	55.934685	-4.404426		M06	55.93979	-4.406
	L07	55.934946	-4.40405		M07	55.93996	-4.40621
	L08	55.935257	-4.402741		M08	55.94017	-4.40632
	L09	55.934428	-4.406126		M09	55.94033	-4.40653
	L10	55.9342	-4.4048		M10	55.94022	-4.40679
	L11	55.934046	-4.405585		M11	55.94012	-4.40672
	L12	55.934982	-4.403692		M12	55.94019	-4.40704
CochnoHigh	H01	55.953987	-4.403044				
	H02	55.95366	-4.40479				
	H03	55.955448	-4.403454				
	H04	55.95296	-4.402511				
	H05	55.9534	-4.404925				
	H06	55.952014	-4.403364				
	H07	55.951696	-4.405688				
	H08	55.95395	-4.405355				
	H09	55.954875	-4.404714				
	H10	55.95533	-4.402822				
	H11	55.955754	-4.404694				
	H12	55.95326	-4.40571				

Table B.2: Sampling locations of Dumgoyne

Field	Samples	Lat	Lon
DumgoyneNorth	N01	56.014566	-4.368024
	N02	56.015517	-4.365849
	N03	56.016559	-4.369705
	N04	56.018919	-4.369283
	N05	56.01513	-4.367409
	N06	56.017999	-4.369136
	N07	56.015049	-4.364936
	N08	56.014996	-4.368815
	N09	56.016775	-4.367557
	N10	56.019737	-4.369765
	N11	56.018416	-4.36791
	N12	56.017648	-4.366292
DumgoyneSouth	S01	56.011444	-4.364544
	S02	56.012303	-4.367148
	S03	56.011552	-4.367088
	S04	56.010912	-4.366178
	S13	56.01191	-4.36758
	S06	56.011375	-4.363492
	S07	56.011783	-4.363527
	S08	56.011424	-4.362811
	S09	56.011647	-4.362803
	S14	56.01088	-4.36521
	S11	56.012786	-4.366714
	S15	56.01208	-4.36325

Appendix C

F. hepatica and environmental factors of each sampling point

Table C.1: CochnoLow data

Samples	Metacercaria	Yield	Yield2	GrassHeight	GrassWeight	SoilTemperature
L01	71.429	0.388	0.039	12.333	184	14
L02	0	0	0	18.500	611	13
L03	42.857	0.482	0.048	7.500	89	14.333
L04	28.571	0.122	0.012	7.833	235	13.233
L05	28.571	0.083	0.008	11.333	343	13.4
L06	57.143	0.295	0.029	13.667	194	14.7
L07	14.286	0.046	0.005	13.333	312	14.767
L08	0	0	0	9.500	212	14.733
L09	28.571	0.102	0.010	8.833	279	13.3
L10	0	0	0	11.333	169	13.3
L11	14.286	0.036	0.004	14.333	396	13.3
L12	28.571	0.077	0.008	14	373	14.733
	SoilMoisture	NDWI	NDVI	Elevation	Slope	Aspect
L01	0.689	0.283	0.560	120.390	1.851	SE
L02	0.622	0.415	0.825	122.950	2.984	SE
L03	0.691	0.270	0.662	120.790	1.674	SW
L04	0.654	0.417	0.823	122.340	2.178	SE
L05	0.577	0.387	0.790	120.720	3.727	SE
L06	0.683	0.201	0.455	120.740	1.235	SE
L07	0.478	0.246	0.492	122.930	2.723	SW
L08	0.570	0.409	0.831	123.890	2.723	SE
L09	0.593	0.440	0.842	123.480	4.151	SE
L10	0.624	0.367	0.758	118.240	0.655	SW
L11	0.647	0.376	0.779	120.040	1.063	SE
L12	0.465	0.322	0.724	122.830	1.221	SE

Table C.2: CochnoMid data

Samples	Metacercaria	Yield	Yield2	GrassHeight	GrassWeight	SoilTemperature
M01	42.857	0.090	0.009	11.667	477	16.833
M02	42.857	0.139	0.014	6.167	308	17
M03	28.571	0.097	0.01	7	296	15.767
M04	28.571	0.129	0.013	5.833	222	15.2
M05	100	0.259	0.026	13.5	386	15.3
M06	57.143	0.147	0.015	12.333	388	15.5
M07	28.571	0.107	0.011	7	267	15.867
M08	42.857	0.177	0.018	7.167	242	15.967
M09	57.143	0.202	0.020	10.5	283	15.9
M10	314.286	1.31	0.131	7.667	240	15.933
M11	14.286	0.04	0.004	15.833	360	15.7
M12	14.286	0.047	0.005	11	301	15.433
	SoilMoisture	NDWI	NDVI	Elevation	Slope	Aspect
M01	0.576	0.404	0.53	172.66	9.677	SE
M02	0.656	0.424	0.604	169.34	2.225	SE
M03	0.657	0.422	0.60	166.74	3.004	SE
M04	0.628	0.409	0.606	164.500	3.195	SE
M05	0.696	0.236	0.542	161.99	2.057	NE
M06	0.676	0.388	0.564	163.97	3.148	SE
M07	0.681	0.421	0.584	165.47	1.99	SW
M08	0.666	0.435	0.62	167.46	2.926	SE
M09	0.625	0.388	0.494	169.32	3.133	SE
M10	0.549	0.371	0.487	168.23	2.015	SW
M11	0.613	0.371	0.599	167.21	2.185	SW
M12	0.435	0.358	0.551	166.94	6.068	SW

Table C.3: CochnoHigh Data

Samples	Metacercaria	Yield	Yield2	GrassHeight	GrassWeight	SoilTemperature
H01	200	0.580	0.058	20.833	345	12.167
H02	0	0	0	18.833	439	13.333
H03	14.286	0.028	0.003	21.500	518	12.733
H04	14.286	0.017	0.002	20.833	820	12.533
H05	57.143	0.119	0.012	18.167	481	13.6
H06	0	0	0	11.667	415	12.2
H07	0	0	0	11.167	184	13.8
H08	14.286	0.024	0.002	18	594	13.3
H09	14.286	0.068	0.007	18.500	210	12.633
H10	0	0	0	17.333	307	12.533
H11	28.571	0.170	0.017	19.167	168	12.267
H12	28.571	0.057	0.006	16	503	12.967
	SoilMoisture	NDWI	NDVI	Elevation	Slope	Aspect
H01	0.744	0.287	0.414	293.670	10.415	NW
H02	0.503	0.416	0.499	283.800	5.226	SE
H03	0.493	0.446	0.449	288.730	2.521	SW
H04	0.489	0.230	0.267	290.610	2.367	SE
H05	0.442	0.396	0.503	283.260	3.488	SE
H06	0.439	0.240	0.201	282.490	4.150	SW
H07	0.555	0.252	0.212	279.070	5.325	SE
H08	0.449	0.428	0.500	288.730	2.366	SE
H09	0.527	0.402	0.421	288.700	1.183	SW
H10	0.594	0.447	0.510	298.580	9.408	SW
H11	0.470	0.459	0.445	289.720	1.543	SE
H12	0.461	0.311	0.477	287.630	3.489	SE

Table C.4: DumgoyneNorth Data

Samples	Metacercaria	Yield	Yield2	GrassHeight	GrassWeight	SoilTemperature
N01	14.286	0.020	0.002	24.750	732	15.133
N02	0	0	0	15.167	401	14.9
N03	14.286	0.020	0.002	25.333	711	14.6
N04	42.857	0.064	0.006	17.833	666	14.567
N05	14.286	0.024	0.002	21.667	592	17
N06	28.571	0.048	0.005	19.833	600	15.7334
N07	14.286	0.018	0.002	23	814	
N08	28.571	0.038	0.004	24.500	756	14.3
N09	28.571	0.041	0.004	22.333	702	18.533
N10	28.571	0.041	0.004	20	692	14.133
N11	385.714	0.471	0.047	23	819	15.366
N12	14.286	0.029	0.003	23.500	490	18.333
	SoilMoisture	NDWI	NDVI	Elevation	Slope	Aspect
N01	0.561	0.396	0.832	23.550	1.155	W
N02	0.598	0.397	0.827	26.030	1.788	SW
N03	0.673	0.359	0.806	22.600	5.304	SW
N04	0.663	0.370	0.809	24.280	1.705	NE
N05	0.656	0.390	0.831	24.300	1.093	NW
N06	0.586	0.404	0.826	21.880	1.644	SE
N07		0.415	0.832	29.420	4.627	SW
N08	0.630	0.393	0.833	23.520	3.276	SW
N09	0.525	0.402	0.836	25.800	1.734	SW
N10	0.512	0.397	0.839	24.280	1.540	E
N11	0.497	0.413	0.839	24.890	2.637	SE
N12	0.491	0.413	0.821	25.960	1.746	SW

Table C.5: DumgoyneSouth Data

Samples	Metacercaria	Yield	Yield2	GrassHeight	GrassWeight	SoilTemperature
S01	14.286	0.052	0.005	13	274	15.133
S02	14.286	0.047	0.005	12.167	302	15.933
S03	42.857	0.168	0.017	10	255	15.967
S04	85.714	0.339	0.034	12	253	15.5
S06	0	0	0	17.833	326	15.4
S07	28.571	0.087	0.009	12.833	328	15.2
S08	0	0	0	18.333	455	15.367
S09				8.667	169	15.933
S11	14.286	0.024	0.002	17.667	603	15.967
S13	57.143	0.174	0.017	14.333	328	15.7
S14	28.571	0.083	0.008	18.667	346	15.167
S15	14.286	0.035	0.004	13.833	404	15.9
	SoilMoisture	NDWI	NDVI	Elevation	Slope	Aspect
S01	0.591	0.305	0.749	25.210	0.545	SW
S02	0.580	0.388	0.823	24.100	0.837	NW
S03	0.636	0.340	0.741	23.910	0.574	NE
S04	0.694	0.315	0.735	23.940	0.182	SW
S06	0.595	0.337	0.764	26.530	0.935	NE
S07	0.564	0.327	0.755	26.650	0.655	SW
S08	0.444	0.315	0.770	28.100	0.581	NW
S09	0.610	0.332	0.772	28.140	0.454	SW
S11	0.641	0.381	0.800	24.410	1.788	SW
S13	0.670	0.333	0.752	23.760	0.908	NW
S14	0.730	0.266	0.704	24.120	1.455	NW
S15	0.610	0.368	0.763	26.940	1.309	SW

Appendix D

Summary of statistical results in Chapter 4

Table D.1: Shapiro-Wilk normality test results in Chapter4

Variables	W-test	p-value	Distribution
Metacercaria	0.504	8.07E-13	Negative binomial
Yield	0.570	7.08E-12	Negative binomial
NDWI	0.917	0.001	Non-normal
NDVI	0.884	3.65E-05	Non-normal
SoilTemperature	0.959	0.046	Non-normal
SoilMoisture	0.955	0.028	Non-normal
Elevation	0.831	8.80E-07	Non-normal
Slope	0.794	9.69E-08	Non-normal
GrassHeight	0.966	0.094	Normal
GrassWeight	0.927	0.001	Non-normal

Table D.2: Kruskal-Wallis test results in Chapter4

Variables by Fields	Kruskal-Wallis chi-squared	df	p-value
Metacercaria	7.8137	4	0.09864
Yield	11.286	4	0.02353
NDWI	10.111	4	0.0386
NDVI	45.145	4	3.71E-09
Elevation	54.003	4	5.255E-11
SoilTemperature	42.108	4	1.584E-08
GrassWeight	25.945	4	0.00003246

Table D.3: One-way ANOVA test results in Chapter4

One-way ANOVA test of Slope					
	Df	Sum Square	Mean Square	F value	Pr(>F)
Field	4	19.89	4.971	13.31	0.000000119
Residuals	55	20.54	0.374		
One-way ANOVA test of Soil Moisture					
	Df	Sum Square	Mean Square	F value	Pr(>F)
Field	4	0.03229	0.008073	3.736	0.00936
Residuals	54	0.11668	0.002161		
One-way ANOVA test of Grass Height					
	Df	Sum Square	Mean Square	F value	Pr(>F)
Field	4	1102.5	275.63	26.44	2.96E-12
Residuals	55	573.4	10.43		

Appendix E

Google Earth Engine JavaScript Codes

```
//Load study areas
var ResearchFields = ee.FeatureCollection('ft:140VAcJBmfVFWlkW-O2Wdq
MTB51DkJMjWH8KFZaoz')
  .set({
    0: {label: 'Field 1'},
    1: {label: 'Field 2'},
    2: {label: 'Field 3'},
    3: {label: 'Field 4'},
    4: {label: 'Field 5'}
  });

//Load Sentinel-2 images
var S2 = ee.ImageCollection('COPERNICUS/S2')
  .filterDate('2016-01-01', '2017-10-10')
  .filterBounds(ResearchFields)
  .filter(ee.Filter.eq('MGRS_TILE','30VVH'));

// clip study fields
var maskFields = function(image) {
  var Fields = image.clipToCollection(ResearchFields);
  return image.updateMask(Fields);
};
var FieldCollection = S2.map(maskFields);

// cloud function to remove clouds
var cloudfunction_ST2 = function(image){
  //use add the cloud likelihood band to the image
```

```
  var quality = image.select("QA60").unmask();
  //get pixels above the threshold
  var cloud01 = quality.gt(0);
  //create a mask from high likelihood pixels
  var cloudmask = image.mask().and(cloud01.not());
  //mask those pixels from the image
  return image.updateMask(cloudmask);
};
var FieldsClear = FieldCollection.map(cloudfunction_ST2);

//NDVI function
var addNDVI = function(image){
  var ndvi = image.normalizedDifference(['B8', 'B4']).rename('NDVI');
  return image.addBands(ndvi);
};
var S2NDVI = FieldsClear.map(addNDVI);
print(S2NDVI);

//Export NDVI images
Export.table.toDrive(S2NDVI.select('NDVI'), 'NDVI');
var size = S2NDVI.size().getInfo();
for (var i = 0; i < size ; i++) {
    var img = ee.Image(S2NDVI.toList(1, i).get(0));
    Export.image.toDrive({
        image: img.select('NDVI'),
        description: i.toString(),
        region: exportarea1
    });
}

//resampling
var resamplingB8 = function(image){
  var prj = image.select('B8').projection() // UTM
  var resampled = image.select('B8').reduceResolution({
      reducer: ee.Reducer.median()
   })
    .reproject(prj.scale(2,2))
    .rename('B8_20');
```

```
  return image.addBands(resampled);
};
var resampled = FieldsClear.map(resamplingB8);

//NDWI function
var addNDWI = function(image){
var ndwi=image.normalizedDifference(['B8_20','B11']).rename('NDWI');
return image.addBands(ndwi);
};
var S2NDWI = resampled.map(addNDWI);
print(S2NDWI);

//NDWI chart
var chartNDWI = ui.Chart.image.seriesByRegion({
  imageCollection: S2NDWI.select('NDWI'),
  regions: ResearchFields,
  reducer: ee.Reducer.mean(),
  scale: 25
});
chartNDWI.setOptions({
  title: 'NDWI B8_20&B11 2016',
   vAxis: {
    title: 'NDWI'
  },
  lineWidth: 1,
  pointSize: 4,
  series: {
    0: {color: '00ff80'},
    1: {color:'ff0000'},
    2: {color: 'ffff00'},
    3: {color: '0080ff'},
    4: {color: 'd358f7'}
  }
});
print(chartNDWI);

//Count clear NDWI pixels
var countNDWI = ui.Chart.image.seriesByRegion({
```

```
  imageCollection: S2NDWI.select('NDWI'),
  regions: ResearchFields,
  reducer: ee.Reducer.count(),
  scale: 25
});
countNDWI.setOptions({
  title: 'NDWI pixels',
   vAxis: {
    title: 'NDWI'
  },
  series: {
    0: {color: '00ff80'},
    1: {color:'ff0000'},
    2: {color: 'ffff00'},
    3: {color: '0080ff'},
    4: {color: 'd358f7'}
  }
});
print(countNDWI);

//Export NDWI images
Export.table.toDrive(S2NDWI.select('NDWI'), 'NDWI B8_20&B11');
var size = S2NDWI.size().getInfo();
for (var i = 0; i < size ; i++) {
    var img = ee.Image(S2NDWI.toList(1, i).get(0));
    Export.image.toDrive({
        image: img.select('NDWI'),
        description: i.toString(),
        region: exportarea1
    });
}

//Visualize NDWI images
Map.centerObject(ResearchFields,12);
Map.addLayer(S2NDWI.select('NDWI'));
Map.addLayer(S2NDVI.select('NDVI'));
```