Advancing Drought Understanding and Prediction in the Vietnamese Mekong Delta

Keke Zhou

Submitted in fulfilment of the requirements for the Degree of

Doctor of Philosophy

School of Social & Environmental Sustainability

College of Social Science

University of Glasgow

September 2024

Abstract

Drought, one of the most destructive climate-related natural hazards, affects millions of people worldwide and poses substantial challenges in the Vietnamese Mekong Delta (VMD), one of Southeast Asia's largest deltas. In recent decades, particularly during 1991-1994, 1998, 2005, 2010, 2015-2016, and 2019-2020, the VMD has suffered from severe and prolonged droughts that resulted in significant socioeconomic impacts. In this delta, droughts often result in severe clean water shortages and extensive damage to cropland. Despite these profound impacts, the mechanisms driving these droughts, including anomalies in the atmospheric moisture transport and land-atmosphere (LA) interactions, and the prediction of droughts in the VMD, remain underexplored. Addressing these gaps is crucial for enhancing drought preparedness and developing effective drought mitigation strategies.

Accordingly, this thesis aims to achieve three primary objectives: (1) to elucidate the sources of precipitation moisture and identify the dominant factors influencing these sources during drought periods in the VMD; (2) to quantitatively assess the LA interactions in the VMD using advanced deep learning techniques; (3) to develop an accurate deep learning model capable of predicting droughts in the VMD on account of atmospheric conditions from the external precipitation source region. The first two objectives are designed to deepen understanding of the mechanisms and processes driving droughts in the VMD, while the third aims to utilize these insights to provide accurate drought predictions.

To better understand the processes of atmospheric moisture transport, the Water Accounting Model-2layers (WAM-2layers), an Eulerian-based moisture tracking model, was employed to identify the primary moisture sources of precipitation in the VMD from 1980 to 2020. In addition, for the first time, the causal inference algorithms were introduced to analyze the causal relationships among variables involved in moisture transport, specifically, to identify which factor drives the moisture transport process and dominates the amount of tracked moisture. The analysis revealed that: (1) over 60% of precipitation in the VMD originates from external moisture sources (60.4%-93.3%), with local recycling contributing from 1.2% to 27.1%; (2) seasonal shifts in monsoon patterns strongly influence the origins of moisture: during the dry season, the South China Sea (northeast) serves as the dominant source, while the Bay of Bengal (southwest) becomes the primary origin during the wet season; (3) based on the causal inference algorithms, atmospheric humidity and wind speed in the upwind area were identified as the principal factors influencing moisture transport during dry and wet seasons, respectively; (4) large-scale forcings (e.g., El Niño and La Niña) were found to

affect the processes of moisture transport significantly and these effects vary spatially and seasonally across the VMD's precipitationshed; (5) local atmospheric conditions, including atmospheric instability (e.g., convective available potential energy, CAPE) and local atmospheric humidity, also play a crucial role in modulating moisture recycling efficiency.

As for the interactions among LA variables, the Long- and Short-term Time-series Network (LSTNet) was applied to model these dynamics over the VMD. The key findings are as follows: (1) the LSTNet model demonstrated superior performance compared to the traditional regional climate model in simulating key variables (i.e., precipitation, soil moisture, sensible and latent heat) during both dry and wet seasons. It exhibited higher accuracy and lower bias, underscoring its suitability for modeling LA interactions in the VMD; (2) this deep learning model effectively captured the relative importance of key variables within the LA interactions, highlighting the critical roles of soil moisture and sensible heat, particularly during dry periods when their negative anomalies substantially reduce precipitation. For example, anomalies in sensible heat were found to decrease precipitation by up to 20% during dry periods, primarily through interactions with temperature and convective inhibition (CIN). Similarly, soil moisture strongly influences precipitation in both dry and wet periods, with deficits leading to reductions in precipitation of up to 30%; (3) projected declines in soil moisture coupled with increases in sensible heat are expected to exacerbate precipitation deficits under changing climatic conditions. By 2075-2099, a 10% increase in sensible heat could reduce precipitation by 3.76% in dry seasons.

Finally, exploring the utility of atmospheric conditions from external precipitation source regions, the deep neural network, Convolutional Gated Recurrent Unit (ConvGRU) was developed to enhance accuracy in drought prediction. The ConvGRU model exhibited exceptional performance in predicting drought conditions at a 3-month lead time, which successfully predicts approximately 90% of meteorological drought events and about 80% of agricultural drought events, with fewer than 10% false predictions for drought months and events. Furthermore, ConvGRU predicts about 70% and 80% compound dry-hot months and events, respectively. The outstanding performance of the ConvGRU model in drought prediction at the 3-month lead time largely attributed to the delayed impacts of external atmospheric conditions, including specific humidity and U- and V-wind, on the VMD's drought conditions through the water vapor transport process. Incorporating the atmospheric data from these external precipitation source regions significantly improves the ConvGRU model's predictive capability, particularly at the lead time of 3 months.

In summary, this research not only advances the understanding of mechanisms driving drought dynamics including external atmospheric moisture transport and local LA interactions, but also establishes an innovative, effective model for drought prediction. These research developments are vital for improving drought resilience and adaptability in the VMD, and offering substantial benefits for regional drought management strategies.

Acknowledgements

Three years ago, I embarked on my PhD journey filled with anticipation and determination. Throughout the three-year journey, firstly, I'd like to express my sincere gratitude to my supervisor, Dr John, for providing me the opportunity to study at the University of Glasgow and for his invaluable guidance and support throughout my research. I also extend my heartfelt thanks to Professor Fabrice for his contributions to my academic development.

My gratitude goes to the College of Social Science for awarding me the PhD scholarship, which was instrumental in funding my research. I am also thankful for the support provided by the academic, technical, and administrative staff at the School of Social & Environmental Sustainability, whose assistance was crucial to my work.

I owe a deep debt of gratitude to my family, my parents and sisters, whose unwavering support and belief in me have been a constant source of spiritual strength, allowing me to focus wholeheartedly on my research. My friends, Hui Wang, Dianyu, Yuting, Yangzi, and Jiren, have enriched my life in Dumfries immeasurably, I am grateful for their friendship and support.

Lastly, I acknowledge my own efforts and endeavours in completing this thesis. Thank you to everyone who has been a part of this journey.

Author’s Declaration

I hereby declare that, except where specific reference is made to the contributions of others, this thesis is the result of my own original research and has not been submitted for the award of any other degree at the University of Glasgow or any other institution.

Chapters 4, 5, and 6 are presented as reformatted versions of research papers that have been published or submitted to academic journals. In accordance with the requirements from ‘Alternative Format Thesis’ guidelines, these chapters collectively represent a coherent and interrelated body of work.

The work in Chapter 4 of this thesis has been published as follows:

Zhou, K., & Shi, X. (2024). Understanding precipitation moisture sources and their dominant factors during droughts in the Vietnamese Mekong Delta. *Water Resources Research*, *60*(7), e2023WR035920. https://doi.org/10.1029/2023WR035920. As the first author and corresponding author of this paper, I confirm that Chapter 4 was jointly authored with Xiaogang Shi, and my contribution to this paper accounts for 95%.

The work in Chapter 5 of this thesis has been published as follows:

Zhou, K., & Shi, X. (2024). Deep learning-based quantitative analyses of feedback in the land-atmosphere interactions over the Vietnamese Mekong Delta. *Science of The Total Environment*, 175119. https://doi.org/10.1016/j.scitotenv.2024.175119. As the first author and corresponding author of this paper, I confirm that Chapter 5 was jointly authored with Xiaogang Shi, and my contribution to this paper accounts for 90%.

The work in Chapter 6 of this thesis has been submitted to the journal:

Zhou, K., & Shi, X. (2024). Leveraging Atmospheric Conditions to Enhance the Drought Predictability over the Vietnamese Mekong Delta. As the first author and corresponding author of this paper, I confirm that Chapter 6 was jointly authored with Xiaogang Shi, and my contribution to this paper accounts for 95%.

List of Abbreviations

ABias	Absolute Bias
BIAs	Relative Bias
CAPE	Convective Available Potential Energy
CCM	Convergent Cross Mapping
CIN	Convective Inhibition
CMIP	Coupled Model Intercomparison Project
CNN	Convolutional Neural Network
ConvGRU	Convolutional Gated Recurrent Units
ECMWF	European Centre for Medium-Range Weather Forecasts
EM-DAT	Emergency Events Database
ENSO	El Niño-Southern Oscillation
ESA CCI	European Space Agency Climate Change Initiative
FAR	False Alarm Rate
FC	Fully Connected
FLEXPART	Flexible Particle
GDP	Gross Domestic Product
GLACE	Global Land Atmosphere Coupling Experiment
GLACE-CMIP5	Global Land Atmosphere Coupling Experiment-Coupled Model Intercomparison Project Phase 5
GNN	Graph Neural Network

GRU	Gated Recurrent Units
IPCC	Intergovernmental Panel on Climate Change
LA	Land-Atmosphere
LSTNet	Long- and Short-term Time-series Network
MCI	Momentary Conditional Independence
MLP	Multilayer Perceptron
NMME	North American Multi-Model Ensemble
PC	Peter and Clark
PCC	Pearson Correlation Coefficient
PCMCI+	Peter and Clark Momentary Conditional Independence plus
PDSI	Palmer Drought Severity Index
POD	Probability of Detection
R^2	Coefficient of Determination
RegCM	Regional Climate Model
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
SPI	Standardized Precipitation Index
SPTI	Standardized Precipitation Temperature Index
SST	Sea Surface Temperature
STI	Standardized Temperature Index
TCW	Total Column Water

VMD	Vietnamese Mekong Delta
VPD	Vapor Pressure Deficit
WAM-2layers	Water Accounting Model with two layers

List of Figures

List of Tables

Table of Contents

Chapter 1. Introduction

1.1 Background

Drought is recognized as one of the most far-reaching, destructive and expensive natural hazards on Earth (Cook et al., 2018; Haile et al., 2020; Alahacoon and Edirisinghe, 2022), which has a wide and significant impacts on ecosystems (Breshears et al., 2005; Zhao and Running, 2010; Van der Molen et al., 2011; Müller and Bahn, 2022; Smith et al., 2024), agriculture (Madadgar et al., 2017; Kuwayama et al., 2019; Lu et al., 2020; Vadez et al., 2024), water supply (Rossi and Cancelliere, 2013; MacAllister et al., 2020; Wang T. et al., 2022), and economy (Ding et al., 2011; Shahpari et al., 2022). Drought also has an indirect impact on health problems including wildfires, dust-related disease, water-related disease, and vector-borne disease (Stanke et al., 2013; Ebi and Bowen, 2016; Berman et al., 2017). Understanding the causes, propagation dynamics, and influencing factors of drought is a necessary condition for implementing drought early warning systems and mitigation measures.

1.1.1 The necessity of drought research

Drought occurs in virtually all geographical areas and brings widespread impacts under various spatiotemporal conditions (Sheffield et al., 2009; Mishra and Singh, 2010; Hao et al., 2018a). For example, drought is the most costly and the second deadliest natural hazards in the United States, which resulted in $250 billion in damages and claimed nearly 3,000 lives between 1980 and 2020 (Ault, 2020). In Asia, the drought event that occurred between 2018 and 2020 resulted in substantial economic losses, amounting to $240 million in Yunnan, China, and $840 million in Thailand (Ha et al., 2022). According to the Emergency Events Database (EM-DAT), more than 600 million people have been affected by drought events over the African continent since 1900, causing around 870,000 recorded deaths and $8.5 billion in economic losses (Ayugi et al., 2022; Delforge et al., 2023). The Millennium Drought, recognized as the worst drought on record in Australia, resulted in widespread water restrictions in major cities, increased electricity prices, significant bushfire incidents, and a 1.6% reduction in Australia’s Gross Domestic Product (GDP) between 2002 and 2003 (Horridge et al., 2005; van Dijk et al., 2013). The spatial scope of recorded drought events exceeds 30% of Europe, and the estimated impact of drought on the overall European economy (excluding social and environmental costs) exceeds €100 billion (European

Commission, 2007; Oikonomou et al., 2020), and the economic losses caused by drought are anticipated to rise swiftly in the future (Naumann et al., 2021). Furthermore, according to the Fifth and Sixth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC), there has been a global increase in both the frequency and severity of droughts (Dai, 2011; Hartmann et al., 2013; Seneviratne et al., 2021), including regions like East Asia (Fischer et al., 2011; Wang et al., 2012; Kang et al., 2022), southern Europe (Sousa et al., 2011; Gudmundsson and Seneviratne, 2015), and West Africa (Dai, 2013). Given the extensive and substantial damages caused by droughts and its increasing frequency under the changing climate, the understanding and prediction of droughts are of great significance and importance in the planning and management of water resources, as well as in the enhancement of drought resilience and adaptability.

1.1.2 Drought definitions and concepts

The occurrence and development of droughts are complex, involving the interplay between various climate processes including large-scale atmospheric circulations (Tao et al., 2014; Kingston et al., 2015; Zhuang et al., 2020), as well as the land-atmosphere feedback (Zeng et al., 2019; Zhou S. et al., 2021; Li Q. et al., 2022). Although drought usually occurs naturally, it can also be caused and aggravated by anthropogenic drivers (Diffenbaugh et al., 2015; Gudmundsson and Seneviratne, 2016; AghaKouchak et al., 2021). Due to the highly nonlinear physical processes of drought, it is difficult to have a universally accepted definition (Hayes et al., 2011; Lloyd-Hughes, 2014; Esfahanian et al., 2017; Mukherjee et al., 2018). The World Meteorological Organization (1992) defined the drought as: (1) "Prolonged absence or marked deficiency of precipitation" and (2) "Period of abnormally dry weather sufficiently prolonged for the lack of precipitation to cause a serious hydrological imbalance". Sheffield and Wood (2011) defined drought as "a deficit of water relative to normal conditions". The definition of drought can vary depending on the variable employed to characterize it, leading to a classification of drought definitions into various categories (Mishra and Singh, 2010). For example, droughts are commonly categorized into meteorological, agricultural, hydrological, and socio-economic types.

Meteorological drought occurs when a region experiences a prolonged period of below-average precipitation (Palmer, 1965; Keyantash and Dracup, 2002; Spinoni et al., 2019). Hydrological drought refers to sustained periods of insufficient surface and subsurface water resources, often assessed using streamflow data (Nalbantis and Tsakiris, 2009; Van Loon, 2015). While droughts typically originate from the insufficiency of precipitation, the evolution from meteorological drought to hydrological drought is not instantaneous and involves additional factors such as limited water storage (Van Loon and Van Lanen, 2012). Consequently, not every meteorological drought necessarily results in a hydrological drought. Agricultural drought involves declining soil moisture that leads to reduced plant growth and crop yields, influenced by factors that also affect meteorological and hydrological droughts, as well as difference between actual and potential evapotranspiration (Boken et al., 2005; Liu X. et al., 2016). Socio-economic drought arises when water resource systems fail to meet water demands, linking drought to the economics of water supply and demand (American Meteorological Society, 2004; Mehran et al., 2015; Guo et al., 2019). It occurs when the demand for water as an economic good exceeds the supply due to a weather-related reduction in water availability (Mishra and Singh, 2010).

Additionally, various other types of drought definitions have been formulated to address specific research questions, including flash drought (Yuan et al., 2015), groundwater drought (Van Lanen and Peters, 2000), vegetation drought (Zhou K. et al., 2021), and ecological drought (Crausbay et al., 2017). Flash drought is characterized by its swift onset and intensification, inducing significant impacts on agriculture and vegetation health, and stressing short-term water resources (Christian et al., 2019, 2021). Groundwater drought, characterized by a reduction in groundwater storage and discharge, results from decreased recharge over a prolonged period of time and these droughts are often exacerbated by human activities, such as groundwater abstraction (Van Lanen and Peters, 2000; Thomas et al., 2017; Han et al., 2019). Vegetation drought represents the vegetation stress and mortality caused by traditional meteorological and agricultural droughts (Brown J. F. et al., 2008; Jha et al., 2019). Crausbay et al. (2017) defined the ecological drought as “an episodic deficit in water availability that drives ecosystems beyond thresholds of vulnerability, impacts ecosystem services, and triggers feedbacks in natural and/or human systems”. In addition, AghaKouchak et al. (2021) introduced the concept of anthropogenic drought, which they define as drought events that are initiated or intensified by human activities. This definition of drought helps to better define and describe the complex interplay between natural processes and human-induced impacts.

Meteorological and agricultural droughts are the primary focus of this study due to their direct and significant impacts on ecosystems and agricultural productivity in the VMD (Lavane et al., 2023). These drought types are inherently linked to physical processes such as land-atmosphere (LA) interactions and atmospheric moisture transport. In contrast, hydrological droughts, while critically important due to their role in driving salinity intrusion and affecting water availability in the VMD, are significantly influenced by human interventions such as dam operations and water management policies (Lu et al., 2014; Kantoush et al., 2017). Furthermore, the limited availability of in situ river discharge data in the VMD presents an additional challenge to studying hydrological droughts comprehensively. Similarly, socio-economic droughts, which involve the interplay between water resource availability and human demand, were not a focus of this study due to the limited availability of socio-economic data for the region.

1.1.3 Development of drought indices

To enhance the assessment and characterization of drought, a number of drought indicators have been developed, each with its advantages and limitations (Mishra and Singh, 2010). For example, Palmer Drought Severity Index (PDSI, Palmer, 1965), Standardized Precipitation Index (SPI, Mckee et al., 1993, 1995), Rainfall Anomaly Index (RAI, van Rooy, 1965), and Standardized Precipitation Evapotranspiration Index (SPEI, Vicente-Serrano et al., 2010) have been developed to quantify the severity and duration of meteorological droughts. The SPI, one of the most popular drought indices because of its simplicity and adaptability, can be flexibly used for short-term and long-term drought assessments by defining different temporal scales (Eslamian et al., 2017; Yihdego et al., 2019). However, the reliability of SPI can be compromised by the length of available precipitation data, as differences in shape and scale parameters in the gamma distribution across various periods can significantly affect the results (Wu et al., 2005). Furthermore, the SPI is calculated only with precipitation data, ignoring factors like evapotranspiration and temperature that influence water demand. To address this limitation, Vicente-Serrano et al. (2010) proposed the SPEI, which incorporates both precipitation and potential evapotranspiration into the calculation. The PDSI, another widely utilized drought index, employs a simple hydrological model based on physical mechanisms, accounting for the supply (precipitation) and demand (temperature) of water. Nonetheless, the PDSI also has certain limitations (Mishra and Singh, 2010): (1) it performs poorly in mountainous and snow-covered areas; (2) it misses much information of hydrological processes due to the oversimplified two-layer model; (3) it responds to drought development and mitigation slowly (Hayes et al., 1999). Several modified versions have been developed since the proposal of PDSI, for instance, Karl (1986) developed the Palmer Hydrological Drought Index (PHDI) for water supply monitoring, while Wells et al. (2004) introduced a self-calibrated Palmer Drought Severity Index (sc-PDSI) that improves the spatial comparability of the PDSI and obtains reliable drought assessment results (van der Schrier et al., 2013).

As for the assessment and evaluation of agricultural drought, soil moisture is the most important factor that needs to be considered in the development of agricultural drought indices (Brocca et al., 2010, 2011; Liu X. et al., 2020; Pan et al., 2023). Notable soil moisture-based drought indices include the Standardized Soil Moisture Index (SSI, Hao and AghaKouchak, 2013; AghaKouchak, 2014), the Soil Water Deficit Index (SWDI, Martínez-Fernández et al., 2015; Mishra et al., 2017; Bai et al., 2018; Zhou K. et al., 2021), the Soil Moisture Agricultural Drought Index (SMADI, Sánchez et al., 2016), the Drought Severity Index (DSI, Cammalleri et al., 2016), and the Soil Moisture Anomaly Percentage Index (SMAPI, Liu et al., 2019). Agricultural drought usually follows meteorological drought, indicating that a precipitation deficit accumulated over three to six months can also serve as a reliable indicator of agricultural drought (McKee et al., 1993; Dai et al., 2020). In addition, research has demonstrated that vegetation conditions can also be an agricultural drought indicator, such as, the Crop Moisture Index (CMI, Palmer, 1968), the Vegetation Condition Index (VCI, Kogan, 1995; Liu and Kogan, 1996; Unganai and Kogan, 1998), and the Vegetation Health Index (VHI, Kogan, 2002; Rojas et al., 2011; Zeng et al., 2023).

The Surface Water Supply Index (SWSI, Shafer and Dezman, 1982), the Standardized Runoff Index (SRI, Shukla and Wood, 2008), and the Streamflow Drought Index (SDI, Nalbantis, 2008; Nalbantis and Tsakiris, 2009) are commonly used hydrological drought indices. The SWSI evaluates surface water availability by integrating snowpack, streamflow, precipitation, and reservoir storage data, thus enhancing the PDSI framework (Shafer and Dezman, 1982, Heim, 2002). The SRI extend the principles of the SPI to monthly runoff data, employing various distributions such as Pearson Type III (PIII), log-logistic, and log-normal for the data fitting (Shukla and Wood, 2008; Peña-Gallardo et al., 2019; Wu, J. et al., 2024). These indices, designed to characterize hydrological drought, generally require extensive data inputs and substantial computational resources. In contrast, the SDI provides a straightforward and effective measure for hydrological drought assessment (Tabari et al., 2013).

In addition to these classical and traditional drought indices, there are still several emerging drought indices have been developed for different purposes. For instance, the Improved Multivariate Standardized Reliability and Resilience Index (IMSRRI, Guo et al., 2019) and the Standardized Water Supply and Demand Index (SWSDI, Wang T. et al., 2022) have been introduced for the assessment of socio-economic drought. The Vegetation Drought Response

Index (VegDRI) integrates climate-based drought indices, satellite-derived data, and other biophysical parameters using data mining techniques to assess drought-related stresses on vegetation (Brown J. F. et al., 2008). Keyantash and Dracup (2004) developed an Aggregate Drought Index (ADI), which comprehensively considers all physical forms of drought by integrating the meteorological, agricultural and hydrological drought indicators through the principal component analysis (PCA). The ADI has not been widely operationalized due to its high data demands and complexity (Tabari et al., 2013). Similarly, the Multivariate Standardized Drought Index (MSDI) probabilistically combines the SPI and the SSI to provide an overall characterization of drought conditions (Hao and AghaKouchak, 2013).

1.1.4 Datasets for drought monitoring and assessment

Accurate and reliable data are essential for drought assessment studies, as well as for managing water resources and forecasting weather (Jiang et al., 2012; Srivastava, 2017). The primary datasets utilized in drought research can be categorized into three types: (1) in-situ measurements, (2) remotely sensed satellite products, and (3) reanalysis datasets (Sun et al., 2018; Tavakol et al., 2021). Among these datasets, the in-situ measurements are often considered the most accurate for monitoring droughts (Zhao et al., 2017). Various large-scale climate datasets have been developed based on these in-situ observations. One such example is the Global Historical Climatology Network daily (GHCNd), which integrates daily climate variables from over 100,000 land surface stations across 180 countries and territories, providing numerous daily variables such as maximum and minimum temperatures and total daily precipitation (Durre et al., 2008, 2010; Menne et al., 2012; Cheng et al., 2024). Another significant resource is the International Soil Moisture Network (ISMN), a comprehensive global in-situ soil moisture database initiated in 2009, which supports the validation and calibration of model- and satellite-derived data and facilitates spatiotemporal analyses (Dorigo et al., 2011a, 2011b, 2013; Gibon et al., 2024). As of 2024, the ISMN has recorded over 11,000 soil moisture time series at over 3,000 stations across more than 30 countries and territories (https://ismn.earth/en/). However, in-situ datasets are limited by their spatial resolution, especially in oceanic and sparsely populated areas (Rana et al., 2015; Kidd et al., 2017). Therefore, to overcome this limitation, several gridded datasets based entirely on in-situ measurements have been developed and are extensively utilized, such as the Climate Research Unit (CRU) Time Series (TS) dataset (Harris et al., 2014), which includes a variety of climate variables like precipitation, mean temperature, and vapor pressure.

Since the launch of the first Television and Infrared Observation Satellite (TIROS) in 1960, (Kidd, 2001), the number of satellite sensors for global atmospheric measurements has expanded significantly (Sun et al., 2018). The rapid development of remote sensing technology enables the measurement and monitoring of key drought-related variables on larger spatial and temporal scales than conventional methods (Choi et al., 2013; Sur et al., 2015). For the measurement of precipitation, extensively utilized satellite-based precipitation products include the Tropical Rainfall Measuring Mission (TRMM, Huffman et al., 2007), the Global Precipitation Measurement (GPM, Hou et al., 2008, 2014; Tapiador et al., 2012), and the Climate Prediction Center (CPC) morphing technique (CMORPH, Joyce et al., 2004). As for the observations of soil moisture based on remote sensing technology, the Soil Moisture and Ocean Salinity (SMOS) mission (Kerr et al., 2001, 2010), the Soil Moisture Active Passive (SMAP, Entekhabi et al., 2010), the European Space Agency climate change initiative (ESA CCI, Hollmann et al., 2013; Dorigo et al., 2017; Gruber et al., 2019), the Sentinel-1 (Torres et al., 2012) and Sentinel-2 (Drusch et al., 2012; Pahlevan et al., 2019) contribute critical soil moisture data for drought research. Satellite-based products also measure evapotranspiration, exemplified by the Global Land Evaporation Amsterdam Model (GLEAM, Miralles et al., 2011; Rouholahnejad Freund et al., 2020), and monitor vegetation conditions through platforms like Landsat (Wulder et al., 2012; Okujeni et al., 2024) and the Moderate Resolution Imaging Spectroradiometer (MODIS, Justice et al., 2002; Gu et al., 2008; Zhang, L. et al., 2017).

As for the reanalysis datasets, the critical idea is to integrate observations with models that encompass a range of physical and dynamical processes, thereby providing a synthesized estimate of the atmospheric state across a uniform grid. This approach ensures spatial homogeneity, temporal continuity, and a multidimensional hierarchy of data (Bosilovich et al., 2008; Sun et al., 2018; Slivinski et al., 2019). A reanalysis system consists of a background forecast model and a data assimilation scheme. The forecast model extrapolates the previous atmospheric state forward in time and space, while the data assimilation scheme merges these forecast outputs with input observations to refine the subsequent atmospheric state (Compo et al., 2011; Fujiwara et al., 2017). The quality of reanalysis products developed by various organizations is gradually improving through the upgrade of modeling approaches, input data, and assimilation techniques. Notable reanalysis products include the European Centre for Medium-Range Weather Forecasts Reanalysis version 5 (ERA5, Hersbach et al., 2020), the Modern Era Retrospective-analysis for Research and Applications (MERRA, Bosilovich et al., 2008; Rienecker et al., 2008; Gelaro et al., 2017), the National Centers for Environmental

Prediction (NCEP) Climate Forecast System Reanalysis dataset (CFSR, Saha et al., 2010), the Twentieth Century Reanalysis (20CR, Compo et al., 2011), and the Global Land Data Assimilation System (GLDAS, Rodell et al., 2004; Liu et al., 2021).

Each type of dataset has its strengths and limitations. For instance, as previously noted, in-situ measurements are often regarded as the most accurate for monitoring droughts, yet they are constrained by limited spatial resolution, particularly in oceanic and sparsely populated regions. On the other hand, remote sensing data, while offering broader coverage, can be compromised by cloud cover, which may affect data quality and introduce uncertainty (Hilker et al., 2012; Prudente et al., 2020). Extensive research has been conducted to evaluate and validate the accuracy and applicability of these datasets in drought assessment and monitoring (Dorigo et al., 2015; Sahoo et al., 2015; An et al., 2016; Katiraie-Boroujerdy et al., 2016; Sun et al., 2016; Colliander et al., 2017; Brito et al., 2021; Jiang et al., 2021; Rahman et al., 2021).

1.2 Droughts in the Vietnamese Mekong Delta

The Mekong River is one of the most important rivers in the world. With its headwaters in the Tibetan Plateau, the river runs through Southwest China (where it is officially called the Lancang River), Myanmar, Laos, Thailand, Cambodia, and southern Vietnam and ultimately discharges into the South China Sea (Kingston et al., 2011; Hecht et al., 2019; Wang J. et al., 2022). The Vietnamese Mekong Delta (VMD), located at the southernmost tip of Vietnam (8.56°-11.03°N, 104.44°-106.82°E, Figure 1.1), marks the endpoint of the Mekong River Basin (Phan et al., 2020). Encompassing 13 provinces, the VMD is bordered by Cambodia to the north and northwest, the Gulf of Thailand to the west and southwest, and the South China Sea to the east and southeast (Phan et al., 2020; Lavane et al., 2023). Spanning approximately 4 million hectares of land area and home to 18 million inhabitants, the VMD contributed up to 18% of Vietnam’s GDP in 2018 (Tran et al., 2019). In the meantime, the VMD plays an important role in the agricultural sector and food security of Vietnam, which supplies half of the nation’s agricultural produce, predominantly rice, vegetables, and aquaculture products (Sebesvari et al., 2012; Tran et al., 2019; Loc et al., 2021; Lavane et al., 2023).

Situated in the tropical monsoon region, the VMD experiences the northeast monsoon during the dry season and the southwest monsoon during the wet season (Vu et al., 2018). In this region, the monthly temperature ranges from 26.1 ℃ to 29.0 ℃, with an average annual precipitation of approximately 2,000 mm (Figure 1.1, calculated based on the ERA5 data from 1980 to 2020). The average precipitation during the dry season (December to April) totals about 250 mm, in contrast to nearly 1400 mm in the wet season (June to October). Additionally, the region records between 2000 and 2500 sunshine hours annually, with higher frequencies occurring in the dry season (Phan et al., 2020).

Over recent decades, the Mekong River Basin has experienced a series of significant drought events, including those during 1991-1994, 1998, 2003-2005, 2010, 2015-2016, and 2019-2020, each resulting in substantial socio-economic impacts (Guo et al., 2017; Lee and Dang, 2018; Kang and Sridhar, 2021; Keovilignavong et al., 2021). Notably, the severe drought event recorded in 2003-2005 (Adamson, 2005; Ha et al., 2023) affected at least 10,000 hectares of winter rice in the VMD, incurring costs of approximately $60 million (Son et al., 2012). The 2015-2016 and 2019-2020 drought periods were particularly catastrophic for the delta. For instance, the 2015-2016 drought resulted in damages exceeding $300 million to agriculture and aquaculture (Nguyen, 2017). In this tidal delta, seasonal seawater intrusion, exacerbated by drought conditions, leads to extensive crop destruction and poses severe threats to the ecosystem and livelihoods of local communities. Specifically, the 2015-2016 drought led to severe salinity intrusion, affecting over one million inhabitants with a clean water shortage and damaging more than 300,000 hectares of farmland (United Nations Resident Coordinator in Viet Nam, 2016; Nguyen, 2017). Moreover, during the 2019-2020 drought, saltwater intrusion reached levels even higher than those observed in the 2015-2016 drought, impacting 82,000 households, exposing more vulnerable populations in VMD to significant water shortage risks, and resulting in estimated agricultural losses of about 460,000 hectares (United Nations Resident Coordinator in Viet Nam, 2020; International Federation of Red Cross Red Crescent Societies, 2020).

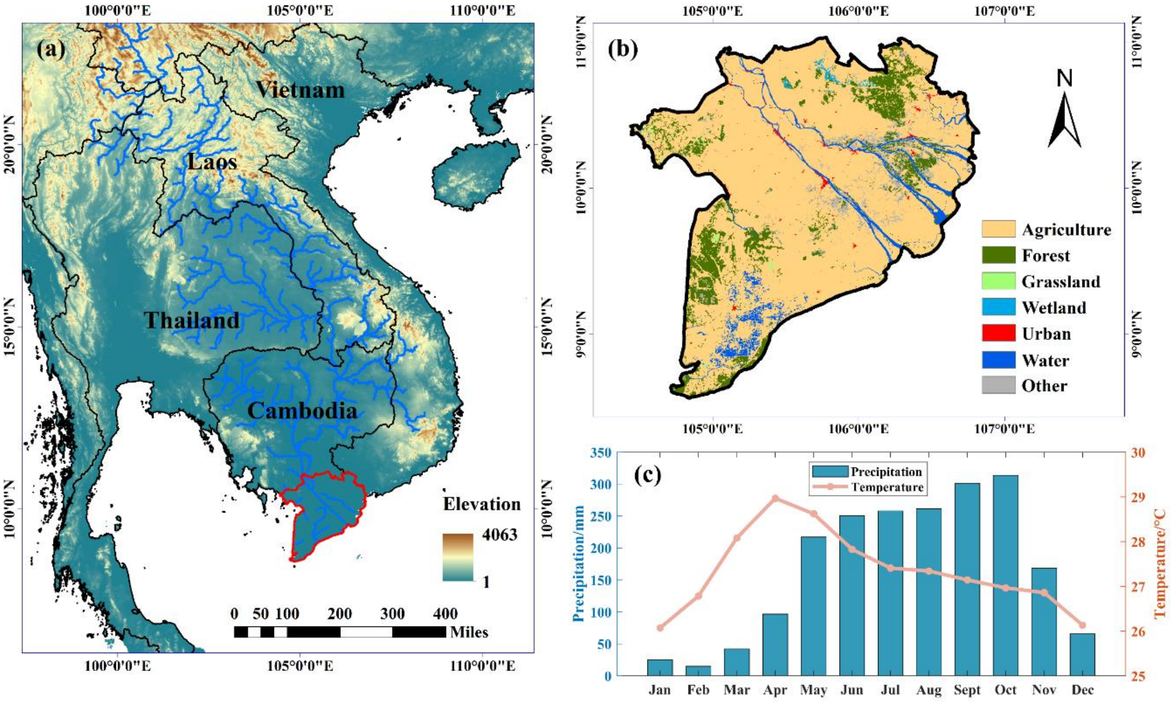

Figure 1.1. The (a) geographic location of Vietnamese Mekong Delta (VMD, within the red boundary), (b) land cover types derived from the European Space Agency Climate Change Initiative land cover datasets, and (c) precipitation and temperature climatology of VMD.

1.3 Drought research over the VMD

Given the crucial role of VMD in Vietnam's agricultural and economic sectors, the severe impact of drought has made it a critical research area that warrants more investigations in recent years in the VMD. Notable studies by Guo et al. (2017), Lee and Dang (2018), and Kang and Sridhar (2021) have analyzed historical droughts using various drought indices. Projections by Sam et al. (2019), Li Y. et al. (2021), Dong et al. (2022), and Kang et al. (2022) suggest that drought frequency, duration, and severity in the Mekong River Basin are likely to intensify in the near future. The study on the drivers of drought in this region points out multiple causes. For example, Lu et al. (2014) and Kantoush et al. (2017) identified dam construction in the upper Mekong River Basin as a factor in reducing downstream discharge and leading to droughts. Cosslett and Cosslett (2018) noted that El Niño conditions typically bring below-average rainfall and shorter wet seasons, contributing to the severe droughts of 2010 and 2015 in the VMD. Furthermore, Bastakoti et al. (2014) and Miyan (2015) attributed the occurrence of droughts to high temperatures and low precipitation.

Although drought causation is complex, it is generally accepted that droughts are associated with large-scale atmospheric circulations (e.g., advection) and terrestrial processes (e.g., LA interactions). The large-scale atmospheric circulation is critical for the initiation of droughts by regulating global moisture transport (Seneviratne et al., 2010; Huang et al., 2024). This insight has aroused interest in the roles of moisture recycling and transport (advection) in drought characterization (Roy et al., 2019). For example, Stojanovic et al. (2021) employed the FLEXPART (FLEXible PARTicle) model to identify moisture sources across seven climate subregions of Vietnam, discovering that precipitation moisture in the VMD predominantly originates from the China Seas during the dry season and the Bay of Bengal during the wet season, with significant correlations observed between contribution moisture anomalies and drought conditions. However, traditional correlation and regression analyses are insufficient for addressing cause-and-effect relationships within the hydrometeorology system (Runge et al., 2019a; Ombadi et al., 2020). The primary factors that drive moisture transport and dominate the VMD precipitation still need to be fully investigated. On the other hand, regarding the LA interactions, they play a key role in the evolution and intensification of droughts (Seneviratne et al., 2010; Seager and Hoerling, 2014; Miralles et al., 2019; Holgate et al., 2020). Previous studies in terms of LA interactions have primarily focused on the water cycle feedbacks, involving soil moisture, evaporation, and precipitation: decreases in soil moisture typically lead to reduced evaporation, drier air, and potentially inhibit the formation of precipitation, thereby increasing the likelihood of drought occurrences (Santanello et al., 2013; Zhou S. et al., 2019; Schumacher et al., 2022). Dirmeyer et al. (2021) further noted that dry soil could alter surface fluxes, dry the atmosphere, and exacerbate droughts and heatwaves over Northern Europe. However, it remains unclear to what extent these factors influence the severity and dynamics of droughts in the VMD. Given the critical roles of moisture transport and land-atmosphere interactions, in light of the understanding of these processes, developing accurate drought prediction models is of great significance for implementing drought early warning systems, as well as for the enhancement of drought resilience and adaptability of VMD.

1.4 Research aims, objectives and questions

This work aims to develop a deep learning model that effectively predicts drought conditions in the VMD based on the understanding of atmospheric moisture transport process and local LA interactions. This is vital and valuable to strengthen drought preparedness and resilience in the region and aid decision-makers in making effective drought management and mitigation strategies. The study is structured around a series of objectives and research questions:

Objective 1: To elucidate the source regions of precipitation moisture and identify the dominant factors influencing these processes during drought periods in the VMD, as detailed in Chapter 4. Research questions include:

- What are the characteristics of the moisture source regions affecting the VMD precipitation?
- Which factors predominantly influence moisture transport and how do they affect the occurrence and evolution of droughts in the VMD?

Objective 2: To quantitatively assess the LA interactions in the VMD using deep learning techniques, as described in Chapter 5. Research questions focus on:

- Can deep learning algorithms effectively capture and simulate the LA interactions in the VMD?
- What is the extent of the influence exerted by key variables within the LA interactions?

Objective 3: To develop a deep learning model that can accurately predict drought conditions in the VMD, as presented in Chapter 6. Research questions are:

- How effectively can the deep learning algorithm predict drought conditions in the VMD?
- Do atmospheric conditions from external precipitation source regions enhance the performance of the drought prediction model?

1.5 Thesis structure

This thesis is structed into 7 chapters to address the objectives and research questions above. Below is an overview of each chapter:

Chapter 1: Provides an introductory background, outlining the motivations behind the study and the broader context of drought research. This includes definitions, classifications, development of drought indices, and necessity for drought research. This is followed by the drought events and studies in the study area, i.e., VMD, along with a brief introduction of existing challenges and research gaps.

Chapter 2: Reviews the current understanding of drought processes and prediction methods, highlighting specific research gaps that warrant further investigation.

Chapter 3: Describes the methodologies employed in this thesis, including moisture tracking models, causality inference algorithms, and deep learning techniques.

Chapter 4: Focuses on the research questions associated with Objective 1. It characterizes the moisture source regions for VMD precipitation, explores the role of upwind moisture in drought propagation, and identifies the key factors influencing moisture transport and VMD precipitation.

Chapter 5: Addresses the research questions for Objective 2 by applying a deep learning algorithm to simulate LA interactions within the VMD and quantitatively assesses the impacts of key variables involved in LA interactions.

Chapter 6: This chapter addressed the research questions of Objective 3, in which a deep neural network was developed to predict drought conditions in the VMD by considering the atmospheric conditions from the external precipitation source region.

Chapter 7: Concludes the thesis by discussing its limitations, contributions to the scientific community, and potential future research directions. It summarizes the main findings, emphasizing their implications and advancements in the drought prediction research.

Chapter 2. Literature Review

This chapter aims to provide a general overview of the research background in terms of drought processes understanding and drought prediction, as well as identifying the research gaps for further investigation. Specific literature reviews related to the research questions are provided in Chapters 4 to 6.

2.1 Understanding how external atmospheric conditions and land-atmosphere interactions affect droughts in the VMD

Precipitation over continental areas is primarily derived from two sources: advection of moisture from regions external to the continent and evapotranspiration from the land surface within the region (Brubaker et al., 1993). Although only about 10% of ocean-evaporated moisture contributes to continental precipitation (Trenberth et al., 2011), it is still the predominant source of precipitation over continents (Sorí et al., 2023). It has been estimated that 36% of global rainfall over land originates from terrestrial evaporation (Van der Ent et al., 2014). Part of land surface evaporation is recycled locally, while the other part is transported hundreds to thousands of kilometres before precipitating (Van der Ent and Savenije, 2011; Wei and Dirmeyer, 2019). Furthermore, deficits in moisture transported from these primary sources are directly linked to drought occurrences (Gimeno-Sotelo et al., 2024). On the other hand, LA interactions play a critical role in the evolution and intensification of droughts (Seneviratne et al., 2010; Seager and Hoerling, 2014; Miralles et al., 2019; Holgate et al., 2020) because evaporation and transpiration from the land surface are also important precipitation sources, which are usually modulated by the LA interactions. For example, reductions in soil moisture generally result in decreased evaporation rates, leading to drier air conditions and potentially suppressing precipitation formation, thereby increasing the likelihood of drought occurrences (Santanello et al., 2013; Zhou S. et al., 2019; Schumacher et al., 2022). Thus, these two processes (i.e., moisture transport and LA interactions) are critical for understanding the mechanisms of drought occurrence and development, enhancing drought predictability in regions like the VMD. The subsequent sub-sections provide a review of studies in terms of these two pivotal processes.

2.1.1 Background of moisture transport studies

Atmospheric moisture transport, a vital element of the hydrological cycle, bridges the major terrestrial and aquatic reservoirs of oceans, lakes, soil, inland and sea ice, and rivers through processes of evaporation, transpiration and precipitation (Brubaker et al., 1993; Savenije, 1996; Van der Ent et al., 2010; Gimeno et al., 2012). Typically, in oceanic regions where evaporation rates surpass precipitation, the oceans serve as net contributors to atmospheric moisture. This moisture is then carried by air currents to continents, where it falls as precipitation. Landmasses usually act as net sinks of atmospheric moisture where precipitation exceeds local evapotranspiration rates. This excess precipitative input supplies rivers, groundwater, and other water bodies, which eventually return it to the oceans, thus completing the hydrological cycle (Gimeno et al., 2012). The import of oceanic moisture is the primary source of precipitation on continents (Gimeno et al., 2010, 2012; Van der Ent and Savenije, 2013), and the moisture transport is an essential link between evaporation from ocean surfaces and precipitation over landmasses (Peixóto and Oort, 1992). In-depth investigations into moisture transport processes not only enhance comprehension of current hydrological changes but also lend empirical support to projections of future climatic conditions (Gimeno et al., 2010, 2012, 2013). Furthermore, a more detailed understanding of these moisture dynamics is crucial for improving precipitation predictions, especially in the regions affected by monsoons.

Following global assessments of moisture source-sink relationships (Yanai and Tomita, 1998; Knippertz et al., 2013; Liu B. et al., 2020), researchers have increasingly focused on regional studies. Over the past few decades, numerous studies have explored moisture transport dynamics at regional scales. For example, the Atlantic Ocean has been identified as a significant moisture source for Europe (Drumond et al., 2011; Pinto et al., 2013; Sodemann and Stohl, 2013; Dayan et al., 2015). Additionally, moisture recycling has been highlighted as particularly crucial for precipitation over the Balkan Peninsula, and Central and Eastern Europe (Bisselink and Dolman, 2008; Sodemann and Zubler, 2010; Ciric et al., 2016). Under dry conditions, continental moisture sources become increasingly important in Europe (Gómez-Hernández et al., 2013; Gimeno et al., 2020). The moisture sources for Asia have significant regional differences. Over China, the Indian Ocean (primarily the Bay of Bengal), the western Pacific, and continental parts of Central and Eastern Asia have been recognized as the main contributors to regional precipitation (Wang and Chen, 2012; Wei et al., 2012; Sun and Wang, 2015; Hu et al., 2018; Xiao and Cui, 2021). During summer, the influence of the Indian and East Asian monsoon systems is notable, particularly over Southern and Eastern China (Ding et al., 2009; Wang and Chen, 2012; Sun and Wang, 2015; Baker et al., 2015; Chen et al., 2021). In India, the Indian Ocean emerges as a dominant moisture source during summer, while moisture recycling plays a significant role in winter (Ordóñez et al., 2012; Pathak et al., 2017). In the United States, the Pacific Ocean and the Gulf of Mexico have been identified as the primary moisture sources (Vachon et al., 2010). These studies underscore the complexity of regional moisture dynamics and their critical role in understanding precipitation patterns across various geographic contexts.

A better understanding of the role of anomalies in moisture transport during extreme hydrometeorological events, such as droughts, is also required. Therefore, recent studies have increasingly focused on the role of moisture sources in diagnosing drought occurrences across multiple spatiotemporal scales (Drumond et al., 2019; Herrera-Estrada et al., 2019). For example, the record-breaking drought that affected western and central Europe in 2016 and 2017 was partly due to diminished moisture transport from the Atlantic, particularly affecting the northern part of the region (García-Herrera et al., 2019). Similarly, Drumond et al. (2017) observed a consistent reduction in moisture contributions from the Mediterranean Sea to central-eastern Europe and the eastern Mediterranean during the most severe summer and winter droughts from 1980 to 2012. Severe drought conditions in major river basins in the South Asian region, such as the Indus, Ganges and Brahmaputra, between 1980 and 2017 were directly linked to decreased precipitation contributions from their climatological moisture sources, including regional basins and the Indian Ocean (Arabian Sea, Bay of Bengal) (Sorí et al., 2017). In China, Guan et al. (2022) examined moisture source anomalies during historic severe droughts in the Mid-to-Lower Yangtze River, noting a reduction in moisture transport from the Pacific Ocean and South China Sea, which primarily resulted in a moisture deficit over the region. Similarly, in southeast Australia, reduced oceanic moisture was identified as a predominant factor influencing drought occurrence and intensification (Holgate et al., 2020). In the United States, the severe 2012 drought in the central regions was associated with reduced moisture contributions from local and continental sources, particularly between June and October, as well as diminished moisture from the mid-latitude Pacific and tropical Atlantic (Drumond et al., 2016; Roy et al., 2019). Herrera-Estrada et al. (2019) demonstrated that in North America, decreased moisture transport, exacerbated by dry soil moisture and reduced evapotranspiration in upwind areas, led to intensified agricultural drought conditions in downwind subregions. As for the VMD, research by Stojanovic et al. (2021) highlighted that the predominant moisture for VMD precipitation during the dry season originated from the China Seas. This result underscores the critical role of external atmospheric conditions from these moisture sources regions in influencing drought occurrence and severity in the VMD, highlighting a complex interplay of regional and distant hydroclimatic influences on drought dynamics. While Stojanovic et al. (2021) identified a significant positive correlation between drought conditions and moisture contribution anomalies, such correlation and regression analyses do not suffice to address cause-and-effect relationships within hydrometeorological systems (Runge et al., 2019a; Ombadi et al., 2020). Thus, the complex causal mechanisms underlying variables in the moisture transport process require more comprehensive investigation, particularly in monsoon regions like the VMD, where precipitation is largely influenced by moisture transport.

2.1.2 Background of land-atmosphere interactions

LA interactions occur at the interface between the land surface and the atmospheric boundary layer, playing a pivotal role in regulating global water and energy exchanges (Beamesderfer et al., 2022). These interactions profoundly influence weather and climate patterns across multiple scales (Santanello et al., 2018). For example, LA feedbacks significantly impact global carbon storage dynamics (Green et al., 2019; Humphrey et al., 2021), soil moisture availability (Shi et al, 2013; Vogel et al., 2017), plant photosynthesis and respiration (Arneth et al., 2012), surface water and energy balance (Salvucci and Gentine, 2013; Lansu et al., 2020), cloud formation (Ek and Holtslag, 2004; Siqueira et al., 2009; De Arellano et al., 2012) and convection (Gentine et al., 2013; Guillod et al., 2014). These interactions also affect atmospheric chemistry and air pollution (Janssen et al., 2013) and are integral to predicting future climate trajectories (Davy and Esau, 2016).

The need to better understand these systematic feedbacks to improve process understanding and model performance has driven the study of LA interactions. Various methods have been employed to characterize LA interactions. For example, Findell and Eltahir (2003) developed a framework using a 1-D coupled model to classify LA coupling into regimes by analyzing convective triggering potential and humidity. Concurrently, studies within the Global Land Atmosphere Coupling Experiment (GLACE) framework, such as those by Koster et al. (2006) and Seneviratne et al. (2006), employed data from multiple forecast and regional climate models to categorize climate and soil moisture regimes based on the impact of soil moisture on evapotranspiration variability. Guo et al. (2006), using the same dataset, highlighted that the variations in the surface water evaporation rate, strongly correlated with soil moisture trends, are a major factor affecting LA coupling strength. Ek and Holtslag (2004) investigated the influence of soil moisture on cloud development based on a coupled 1-D land surface-PBL (planetary boundary layer) model, emphasizing the measurements of entrainment fluxes are important for understanding the LA feedbacks. Santanello et al. (2009, 2011, 2013) highlighted the importance of quantifying entrainment during the assessment of LA coupling strength over the southern Great Plains in the United States. In addition, there are several LA coupling metrics, such as the convective triggering potential and low-level humidity index ($CTP\text{-}HI_{low}$, Findell and Elthair, 2003), the mixing diagrams (Santanello et al., 2009), the heated condensation framework (HCF, Tawfik and Dirmeyer, 2014), were utilized to characterize LA feedbacks and LA coupling processes (Santanello et al., 2018; Wakefield et al., 2021).

Observations with extensive spatial and temporal coverage are essential for the evaluation and development of LA models, which typically operate over large ranges (Seneviratne et al., 2010; Guillod et al., 2014). Several regional and global observation networks have been established to support the research on LA interactions, such as the AmeriFlux (Novick et al., 2018), the EuroFlux (Valentini, 2003), the AsiaFlux (Mizoguchi et al., 2009), the Ozflux (Beringer et al., 2016), and the FLUXNET (Baldocchi et al., 2001). These datasets have proven invaluable for improving understanding of how land surface responds to atmospheric forcing over timescales ranging from seconds to years. In addition, several field campaigns and research networks have been developed in recent years to enhance the observation and understanding of LA interactions. Notable initiatives include the Global Energy and Water Exchanges Project (GEWEX, Chahine, 1992), the GEWEX Global Land-Atmosphere System Study (GLASS, van den Hurk et al., 2011), the Land-Atmosphere Feedback Experiment (LAFE, Wulfmeyer et al., 2018), the Local Land–Atmosphere Coupling (LoCo) Working Group (Santanello et al., 2018), the Chequamegon Heterogeneous Ecosystem Energy-Balance Study Enabled by a High-Density Extensive Array of Detectors 2019 (CHEESEHEAD19, Butterworth et al., 2021), and the GEWEX Land-Atmosphere Feedback Observatories (GLAFO, Späth et al., 2023). These efforts are critical for advancing understanding of LA interactions and their implications for climate and weather forecasting.

Soil moisture is a critical component in LA interactions and broader climate systems, significantly influencing drought occurrence and intensification through the water cycle feedbacks (e.g., soil moisture-evaporation-precipitation): the decline of soil moisture typically reduce evaporation, which can dry the atmosphere, and potentially suppress precipitation formation, thereby increasing the likelihood of drought occurrence (Santanello et al., 2013; Zhou S. et al., 2019; Schumacher et al., 2022). For example, Alessi et al. (2022) observed that strengthened LA coupling in the northeastern United States, accompanied by a positive soil moisture-rainfall feedback, is likely to increase drought frequency. Similarly, Wang and Yuan (2022) demonstrated that the drier land surface makes the atmosphere drier through LA coupling, accelerating drought onset in China by reducing precipitation and increasing evapotranspiration. LA interactions not only affect the drought development locally but also impact the drought conditions in downwind areas. For instance, Schumacher et al. (2022) noted that reduced evaporation due to dry soils can affect the land surface energy balance, influencing both local and downwind precipitation. In drylands, precipitation reductions during drought events can exceed 15% per event and up to 30% per month, with these feedback mechanisms potentially intensifying future droughts. Similarly, Zeng and Yuan (2021) pointed out that the LA coupling over the upstream region (i.e., south of Lake Baikal) in an earlier stage may alter atmospheric circulation patterns and affect drought intensity and persistence in downstream region (i.e., Northeast China).

Additionally, the reduction in soil moisture can significantly increase sensible heat and temperatures, potentially triggering heatwaves (Hirschi et al., 2011; Mueller and Seneviratne, 2012; Miralles et al., 2014; Geirinhas et al., 2021). Prior research has extensively documented the concurrent droughts and heatwaves caused by the interplay between soil moisture deficits and elevated temperatures (Hao et al., 2018b; Schumacher et al., 2019). For example, Seo and Ha (2022) observed a decline in soil moisture across northern East Asia since the late 1990s, which led to an increase in evaporative stress and eventually amplified compound drought and heatwaves in the region. Similarly, Dirmeyer et al. (2021) pointed out dry soil would alter surface fluxes, dry atmosphere, and exacerbate the drought and heatwave conditions over Northern Europe. Moreover, accurate understanding and characterization of LA interactions is crucial for enhancing the predictability of drought (Roundy et al., 2014). Nonetheless, to what extent LA variables affect the process and severity of drought still lack research.

Previously, to assess the atmospheric response to variations in land surface conditions, the GLACE was initiated, employing 12 Atmospheric General Circulation Models (AGCMs) to quantify LA coupling strength (Koster et al., 2004, 2006; Guo et al., 2006). Moreover, recognizing the pivotal role of soil moisture in LA interactions, within the framework of the Coupled Model Intercomparison Project Phase 5 (CMIP5), the GLACE-CMIP5 experiment was implemented to evaluate the effects of soil moisture on the long-term changes in climate for both historical and future scenarios (Seneviratne et al., 2013; Schwingshackl et al., 2018). However, while the GLACE-CMIP5 experiment focused on how soil moisture variability influences the climate system, particularly in terms of precipitation and temperature, it overlooked the influence of other critical climate factors such as sensible heat. Consequently, there remains a gap in the quantitative analysis of the interactions within the coupled water and energy balances (soil moisture-sensible heat-precipitation) associated with anomalies in sensible heat and precipitation. A comprehensive exploration of these inter-relationships within LA interactions, and a detailed quantitative analysis of the relative contributions of each variable, could be effectively advanced through the application of deep learning algorithms. This approach would allow for a more nuanced understanding of the complex dynamics governing climate systems.

2.2 Background of drought prediction studies

As illustrated in Chapter 1, droughts have left severe destruction on ecosystems, agriculture and economies worldwide (Smith and Matthews, 2015; Madadgar et al., 2017; Ha et al., 2022; Müller and Bahn, 2022). In light of their extensive damages and increasing frequency under the changing climate, predicting droughts is crucial for effective water resource management and enhancing drought resilience (Fung et al., 2020; Al Mamun et al., 2024; Jadhav et al., 2024). Therefore, a large number of research has been dedicated to drought prediction across various regions including Asia (Rhee and Im, 2017; Xu et al., 2019; Zhang et al., 2019, 2021; Aghelpour et al., 2020, 2021; Malik and Kumar 2020; Li J. et al., 2021a; Malik et al., 2021; Wu et al., 2021; Mohammed et al., 2022; Wu, Z. et al., 2022; Lin et al., 2023; Wang et al., 2023; Al Mamun et al., 2024; Jadhav et al., 2024), Europe (Bonaccorso et al., 2015; Slater et al., 2017; Başakın et al., 2021; Li J. et al., 2021a; Danandeh Mehr et al., 2023), Africa (Djerbouai and Souag-Gamane, 2016; El Ibrahimi and Baali, 2018; Shukla et al., 2019; Achour et al., 2020; Li J. et al., 2021a; Achite et al., 2023a, 2023b), Australia (Schepen et al., 2014; Dikshit et al., 2020a, 2020b), and United States (Bolinger et al., 2017; Yan et al., 2017; Mo and Lettenmaier, 2020; Jiang and Luo, 2022; Cao et al., 2023; Latifoğlu and Özger, 2023). The methods utilized for drought prediction in these studies are primarily categorized into statistical methods, dynamical models, and hybrid approaches (Dikshit et al., 2021a; Nandgude et al., 2023). Statistical models employ a variety of predictors identified from historical hydroclimatic records (including oceanic, atmospheric, and land components), analyzing causal relationships between these variables and different drought indices to predict different types of droughts (Hao et al., 2018a; Xu et al., 2018a; Barrett et al., 2020). Dynamical prediction models, such as the National Centers for Environmental Prediction (NCEP) coupled forecast system model version 2 (CFSv2, Saha et al., 2014; Siegmund et al., 2015), European Centre for Medium-Range Weather Forecasts (Bonavita et al., 2016;

Johnson et al., 2019), simulate real land-atmosphere-ocean physical interactions and processes based on climate and hydrologic models (Hao et al., 2018a; Xu et al., 2018a). In recent years, several multi-model ensembles have been developed and applied in precipitation or drought forecast (Mo and Lyon, 2015; Ma et al., 2016; Xu et al., 2018b), such as the North American Multi-Model Ensemble (NMME, Kirtman et al., 2014). Generally, dynamical models are developed to simulate and forecast weather and climate conditions, with precipitation and temperature predictions being utilized to calculate drought indices (Dutra et al., 2014; Hao et al., 2018a). Hybrid models blend the strengths of both statistical and dynamical approaches, aiming to leverage their respective advantages to enhance drought prediction accuracy (Xu et al., 2018a; Aghakouchak et al., 2022). For example, Yan et al. (2017) developed a probabilistic drought forecasting framework integrating dynamical and statistical models based on a copula function to improve drought forecasting skills. Similarly, Madadgar et al. (2016) combined the NMME model with a Bayesian-based statistical model using the Expert Advice algorithm (Cheng and AghaKouchak, 2015), achieving superior performance compared to the standalone NMME model. These hybrid approaches underscore the potential for refined drought prediction models that effectively integrate diverse modeling strategies.

Statistical methods, due to their simplicity and effectiveness, have been prominently utilized in the aforementioned drought prediction research. Such as, multiple Markov chains were proposed and evaluated by Cao et al. (2023) for categorial drought prediction on the U.S. Drought Monitor at a weekly scale. A distributed lag nonlinear model was developed by Zhang et al. (2019) for meteorological drought forecasting in Shaanxi province, China. Additionally, Wu, H. et al. (2022, 2023) utilized copula-based drought prediction methods to effectively predict seasonal hydrological and agricultural drought, while Bonaccorso et al. (2015) introduced probabilistic models to forecast drought class transitions based on the SPI. In addition to these statistical methods, machine learning algorithms and deep neural networks have also been extensively used in drought prediction due to their efficiency and accuracy. The methods based on Artificial Neural Networks (ANN) have been explored by Belayneh et al. (2016), Tufaner and Özbeyaz (2020) and Banadkooki et al. (2021). Similarly, Random Forest (RF, Park et al., 2016; Mohammed et al., 2022; Al Mamun et al., 2024) and Support Vector Machine (SVM, Belayneh and Adamowski, 2012; Belayneh et al., 2014; Aghelpour et al., 2021; Malik et al., 2021) have been applied for drought monitoring and prediction. The models based on Long Short-Term Memory (LSTM) were also extensively used in drought prediction (Kaur and Sood, 2020; Danandeh Mehr et al., 2023; Wang et al., 2023; Grabar et al., 2024).

Among these statistical predictions of droughts, several studies have highlighted the importance of incorporating large-scale atmospheric forcing into the prediction models. For instance, Zhang et al. (2021) demonstrated that integrating the El Niño-Southern Oscillation (ENSO) into the meta-Gaussian model significantly enhanced the predictability of agricultural drought in regions impacted by large-scale atmospheric circulations. Similarly, Bonaccorso et al. (2015) observed improved forecasting performance when incorporating the North Atlantic Oscillation (NAO) Index. Özger et al. (2012) and Ganguli and Reddy (2014) also found that using lagged climatic variables (e.g., ENSO, the Indian Ocean Dipole Mode, and the Atlantic Multidecadal Oscillation) as input improved forecasting capabilities. However, instead of simply utilizing time-series data of atmospheric circulations as inputs, incorporating detailed spatial-temporal atmospheric conditions may enhance regional drought prediction capabilities. For example, Holgate et al. (2020) found that drought occurrence and intensification in southeast Australia were predominantly influenced by reduced oceanic moisture. Similarly, Stojanovic et al. (2021) noted that precipitation moisture for the VMD primarily originated from the China Seas during the dry season. This result means the atmospheric conditions in the external precipitation source region (e.g., China Seas) of the target area (i.e., VMD) play a critical role in the occurrence and intensification of drought. Thus, incorporating atmospheric conditions in the external area into the prediction models, rather than solely relying on the time series of large-scale forcings as predictors, is likely to improve the performance of deep neural networks in predicting sub-seasonal to seasonal droughts. This approach promises a more nuanced and effective strategy for drought prediction, leveraging advanced deep learning techniques to integrate comprehensive atmospheric data.

Chapter 3. Overview of the methods

Chapters 1 and 2 illustrated the study background, including research aims, objectives and questions, and corresponding literature review. This chapter aims to provide a general overview of methods used in the thesis to achieve the research objectives and bridge the research gaps identified in previous chapters. Refer to Chapters 4 to 6 for specific and detailed information on the methods used in this thesis.

3.1 Moisture tracking models

As outlined in Section 2.1, the external atmospheric conditions from the moisture sources regions and moisture transport processes are pivotal in influencing both the occurrence and severity of droughts (Drumond et al., 2019; García-Herrera et al., 2019; Herrera-Estrada et al., 2019; Guan et al., 2022). This underscores the effects of distant hydroclimatic factors on drought dynamics. Consequently, a reliable and accurate moisture tracking model is crucial for elucidating the mechanisms of moisture transport and identifying the origins of precipitation moisture. To track atmospheric water vapor, several methods are employed: (1) analytical and box models, (2) numerical water vapor tracers, and (3) physical water vapor tracers (Gimeno et al., 2012, 2020).

Analytical and box models utilize the Eulerian framework to analyze the vertically integrated balance of water vapor (Burde and Zangvil, 2001; Gimeno et al., 2020). This approach considers the budget of evapotranspiration minus precipitation by examining changes in the storage of water vapor within the integrated column and the divergence of the integrated vapor transport:

$$E - P = \frac{\partial W}{\partial t} + \nabla \cdot \vec{\Phi} \tag{3.1}$$

where E and P represent evapotranspiration and precipitation; W represents the storage of water vapor in the integrated column; and $\vec{\Phi}$ represents the divergence of the integrated vapor transport. The storage of water vapor (W) and the divergence $\vec{\Phi}$ can be assessed by integrating both specific humidity (q) and the horizontal wind field ($\mathcal{V}$) across the vertical column, extending from surface pressure (Ps) to upper atmospheric levels:

$$W = \frac{1}{g}\int_{P_s}^{0} q\,dp \tag{3.2}$$

$$\vec{\Phi} = \frac{1}{g}\int_{P_s}^{0} q\mathcal{V}dp \tag{3.3}$$

While analytical and box models provide a straightforward framework to evaluate the vertically integrated balance of water vapor, they rely on several assumptions, such as negligible changes in atmospheric water storage or the assumption of a well-mixed atmosphere. Although these assumptions facilitate ease of implementation and reduce computational demands, they limit the models' ability to accurately depict the physical processes in moisture transport (Gimeno et al., 2020).

The study of source-sink regions for atmospheric moisture can also employ numerical water vapor tracers, often referred to as a "water vapor tagging" approach. This methodology is usually classified into Eulerian (Joussaume et al., 1984; Koster et al., 1986; Bosilovich and Schubert, 2002; Rios-Entenza et al., 2014; Insua-Costa and Miguez-Macho, 2018) and Lagrangian (Stohl and James, 2004, 2005; Sun and Wang, 2014; Chu et al., 2017) techniques. Eulerian methods focus on fixed geographical points to monitor the passage of air masses, while Lagrangian techniques track the forward and backward trajectories of air particles, thereby mapping the paths of humid air parcels. Lagrangian methods have gained prominence in recent years for determining the origins of moisture leading to precipitation in specific regions and for analyzing the linkage between atmospheric moisture transport and drought phenomena (Stojanovic et al., 2017, 2018; Drumond et al., 2019; Guan et al., 2022). For example, based on the Lagrangian methods, Dirmeyer et al. (2014) highlighted significant shifts in evaporative moisture sources during droughts in arid and semiarid regions. Utilizing 3D Lagrangian model FLEXPART, Drumond et al. (2017), García-Herrera et al. (2019) and Benedict et al. (2021) demonstrated that severe drought events in Europe correlate significantly with anomalous moisture contributions from the oceans (e.g., North Atlantic Ocean and Mediterranean Sea). Similar phenomena have been observed globally based on the Lagrangian-based models, including Indus, Ganges and Brahmaputra River basins (Sorí et al., 2017), Texas and Upper Midwest Region in United States (Roy et al., 2019), southeast Australia (Holgate et al., 2020), China (Chu et al., 2017; Guan et al., 2022), and VMD (Stojanovic et al.,2021).

In contrast to Lagrangian models, which track air particles along with dynamic humidity data (Tuinenburg and Staal, 2020), the Eulerian framework analyzes fluid movement at fixed spatial coordinates over time (Li Y. et al., 2022). In fact, Lagrangian techniques are preferred for broad climatological studies due to their capacity to track moisture over large scales and timeframes, whereas Eulerian methods are more suited to detailed, location-specific case studies. For example, the Weather Research and Forecast (WRF) Water Vapour Tracer (WRF-WVT) has been effectively applied to assess moisture recycling in the Iberian Peninsula (Rios-Entenza et al., 2014), to explore moisture sources affecting the North American monsoon (Dominguez et al., 2016), and to study precipitation recycling over the Tibetan Plateau (Gao et al., 2020). Additionally, the Water Accounting Model (WAM-2layers) was used by Xiao and Cui (2021) to identify moisture sources for seasonal precipitation in the Pearl River Delta, China, and by Zhang (2020) to investigate moisture sources during extreme droughts in Southwest China.

Furthermore, physical water vapor tracers, specifically stable water isotopes, offer a unique fingerprinting mechanism to trace the origins of atmospheric water (Gimeno et al., 2020). These tracers are invaluable not only in modern climatology but also in the interpretation of paleoclimatic records, providing insights into historical moisture sources and weather patterns (Jouzel et al., 2013).

The Water Accounting Model with two layers (WAM-2layers) utilized in this study is an Eulerian-based framework initially developed by Van der Ent et al. (2010). The WAM model was firstly designed to highlight the significant influences of global wind patterns, topography, and land cover on continental moisture recycling and transport. Originally, the WAM model operated under a “well-mixed assumption” with a single vertical layer, which proved inadequate for accurately reproducing the patterns of evaporation and precipitation (Van der Ent et al., 2013). The number of layers in the vertical and the mixing assumption after evaporation had the largest influence on the moisture transport modelling, especially in contexts of strong wind shear (Van der Ent et al., 2013). The subsequent version, WAM-2layers, addressed these limitations by incorporating two strategically defined vertical layers, enhancing the model’s ability to simulate moisture transport. The refined model, WAM-2layers, has been extensively applied in recent years for tracking evaporative contributions to precipitation on both regional and global scales, demonstrating robust capabilities for rapid computation in large-scale and long-term atmospheric moisture tracking studies (Keys et al., 2012; Keys et al., 2014; Van der Ent, 2014; Van der Ent et al., 2014; Zhang, 2020; Xiao and Cui, 2021; Li Y. et al., 2022). The operational principle of WAM-2layers is grounded in the atmospheric water balance equation, as outlined by Van der Ent et al. (2014):

$$\frac{\partial S_k}{\partial t} = \frac{\partial (S_k u)}{\partial x} + \frac{\partial (S_k v)}{\partial y} + E_k - P_k + \varepsilon_k \pm F_v \tag{3.4}$$

where S_k indicates the atmospheric moisture storage in layer k (the top or bottom layer), E and P indicate the evaporation and precipitation, u and v indicate the wind speed in the zonal- and meridional-direction, respectively, ε indicates the residual, and F_v is the vertical moisture transport; $\frac{\partial (S_k u)}{\partial x}$ and $\frac{\partial (S_k v)}{\partial y}$ represent the horizontal moisture transport. Further details regarding the WAM-2layers model are comprehensively discussed in Chapter 4.

3.2 Causality inference algorithms

Causal inference is the process of determining whether a specific relationship between variables can be described as causal, distinct from mere correlations (Guo et al., 2020). This field lies at the intersection of statistics, philosophy, and computer science, which methodically analyze the effects of actions, interventions, or natural occurrences on outcomes (Nogueira et al., 2022). It is well known that correlation does not imply causation. The causal inference seeks to establish direct influences among variables, rather than simple associations (Yao et al., 2021; Nogueira et al., 2022). The significance of causal inference extends across various scientific domains, including statistics, computer science, education, public policy, economics, and Earth and environmental sciences (Gangl, 2010; Glass et al., 2013; Imbens and Rubin, 2015; Pearl et al., 2016; Varian, 2016; Massmann et al., 2021; Runge et al., 2023).

The formal statistical approach to causation can be traced back to randomized controlled trials in the early 20th century, offering a robust method for inferring causality by mitigating confounding influences through randomization (Rubin, 2005). This methodology laid the groundwork for the Neyman-Rubin Causal Model (NRCM), which introduces the concept of potential outcomes and evaluates the possible outcomes that may arise from different treatment states within the same unit (Rubin, 1974; Holland, 1988; Pearl, 2010). The latter half of the 20th century witnessed significant contributions that shaped the contemporary landscape of causal inference. Notable among these was the Granger causality (GC) test (Granger, 1969), which applied linear vector autoregression to test for causal relationships, and the directed acyclic graph (DAG) framework introduced by Pearl (1995), which allows for the representation of causal relationships among variables (Pearl, 2000). Nowadays, the availability of observational data and advances in computational power have shifted the focus towards estimating causal effects from observational datasets (Runge et al., 2019a; Yao et al., 2021). Specifically, in hydrometeorology research field, the accumulation of many in-situ and remote sensing data has spurred the development of data-driven causal analysis methods (Runge et al., 2019a; Ali et al., 2024). For example, there are several toolbox namely Tetrad (Ramsey et al., 2018), causal-learn (Zheng et al., 2024), and Tigramite (Runge et al., 2019b; Runge, 2020), which incorporates various of causal discovery algorithms including the Granger causality test, the Peter Clark (PC) algorithm (Spirtes and Glymour, 1991), the linear non-Gaussian acyclic mode (LiNGAM, Shimizu et al., 2006), the Peter and Clark Momentary Conditional Independence (PCMCI, Runge et al., 2019b; Runge, 2020). Each method has its own strengths and weaknesses. In this research, two prominent causal inference algorithms, namely the Convergent Cross Mapping (CCM, Sugihara et al., 2012) test and the PCMCI plus (PCMCI+), were integrated to leverage their respective advantages to elucidate causal relations within the atmospheric moisture transport process.

The CCM method was developed to identify causal relationships within dynamical systems to address the limitation of assuming variables are stochastic. It has been applied to assess the influence of soil moisture on precipitation (Wang et al., 2018), and the effects of various environmental factors such as air temperature, vegetation index, soil moisture, net surface radiation, precipitation on land surface temperature (Wu, J. et al., 2023). However, it has been noted by Ombadi et al. (2020) that CCM can produce erroneous bidirectional causality results in cases where variables are strongly coupled, even though the true relationship may be unidirectional. In parallel, the graph-based PC algorithm, conceived by Peter Spirtes and Clark Glymour, stands as another widely used method in causal inference. The PC algorithm was applied to analyze the environmental drivers of evapotranspiration in the research of Ombadi et al. (2020). It has been further developed into the PCMCI and PCMCI+ algorithms by Runge et al. (2019b) and Runge (2020), which further conduct a Momentary Conditional Independence (MCI) test after the PC algorithm. The causality directions of the PCMCI+ algorithm and the causality strength of the CCM were integrated to comprehensively evaluate the causal relations of variables in the moisture transport process. Further details regarding the PCMCI+ and CCM test are comprehensively introduced in Chapter 4.

3.3 Deep learning algorithms

Deep learning, an influential subset of machine learning, has substantially impacted a range of scientific fields (LeCun et al., 2015; Guo et al., 2016; Min et al., 2017; Kamilaris and Prenafeta-Boldú, 2018; Shinde and Shah, 2018), including hydrometeorology (Lee et al., 2021; Zhi et al., 2021; Liu et al., 2022a; Tripathy and Mishra, 2024). The origins of deep learning can be traced back to the foundational period of Artificial Intelligence in the 1950s (Dick, 2019). Significant developments occurred in the late 1990s and early 2000s, with the creation of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to address data with spatial and temporal dependencies, respectively. CNNs, which process data with grid-like topologies such as images, utilize convolution operations to effectively capture spatial structures (LeCun et al., 1998; Krizhevsky et al., 2012). Over the recent years, CNNs have undergone extensive advancements, leading to robust architectures like AlexNet (Krizhevsky et al., 2012), VGGNet (Simonyan and Zisserman, 2015), GoogLeNet (Szegedy et al., 2015), and ResNet (He et al., 2016), which have substantially advanced the fields of image and video recognition. RNNs, suited for sequential data such as text, are designed to account for dependencies where current inputs rely on preceding elements in the sequence. This makes them particularly effective for natural language processing and time-series analysis (Rumelhart et al., 1986; Bengio et al., 1994; Mikolov et al., 2013; Pascanu et al., 2014). However, RNNs often struggle with long-range sequence training due to vanishing or exploding gradients (Hochreiter, 1998; Dang et al., 2017). The LSTM (Hochreiter and Schmidhuber, 1997) and Gated Recurrent Units (GRU, Chung et al., 2014) are two improved models to solve this issue, which selectively retain or exclude information from the sequence based on the gate units. RNN-based models have been widely applied to handle sequential learning problems (Cho et al., 2014; Zhang et al., 2020). Further, there is a widely used deep neural networks namely Convolutional LSTM (Conv-LSTM), which integrates the spatial processing strengths of CNNs with the temporal modeling capabilities of LSTM, making it ideal for applications involving temporal and spatial dimensions of data, such as predicting the movement and intensity of storms from satellite images (Shi et al., 2015, Liu et al., 2017; Mohd Noor et al., 2022). In addition, recent innovations have seen the rise of Graph Neural Networks (GNNs), which model the complex relationships and interactions between various components (Gori et al., 2005; Scarselli et al., 2009; Gallicchio and Micheli, 2010; Dai et al., 2018).

Deep learning algorithms have experienced rapid development and broad adoption within geosciences (Reichstein et al., 2019; Yu and Ma, 2021), contributing significantly to research in natural hazards and extreme climate events (Liu and Wu, 2016; Liu Y. et al., 2016; Sharma et al., 2017; Racah et al., 2017), as well as the advancements in spatial and temporal state prediction, including precipitation nowcasting (Shi et al., 2015; Zaytar and Amrani, 2016;

Zhang et al., 2023). These methods also support complex Earth system modelling (Gao et al., 2022; Bi et al., 2023) and enhance data assimilation, downscaling and blending (Vandal et al., 2018; Niu et al., 2020; Liu et al., 2022b). Trained on extensive and various datasets, these models excel at learning from historical patterns and accurately forecasting future conditions. Their applications are increasingly crucial for real-time decision-making in weather forecasting, long-term climate simulations, and strategic responses to climate-related disasters (Zhou K. et al., 2019; Zhang et al., 2022).

The deep neural networks utilized in this work include: (1) the Long- and Short-term Time-series Network (LSTNet, Lai et al., 2018), which integrates CNNs, RNNs (i.e., GRU) and fully connected (FC) layers. The utilization of LSTNet aims to simulate the interactions among the LA variables because it can extract short-term dependency patterns among the variables with CNNs and discover long-term temporal patterns through the RNN layers; (2) the convolutional GRU (ConvGRU) model, which comprises CNN layers enhanced with spatial-channel wise attention mechanisms (Woo et al., 2018), a GRU layer and a multilayer perceptron (MLP) layer. The ConvGRU is specifically developed to evaluate the performance of deep learning algorithms in predicting drought conditions, taking into account the atmospheric conditions from the external precipitation source region. Further details regarding the LSTNet model can be found in Chapter 5 and Lai et al. (2018), and detailed descriptions of the ConvGRU model are provided in Chapter 6. Overall, Figure 3.1 provides the overview of the methodologies used to achieve the research aims and objectives in this study.

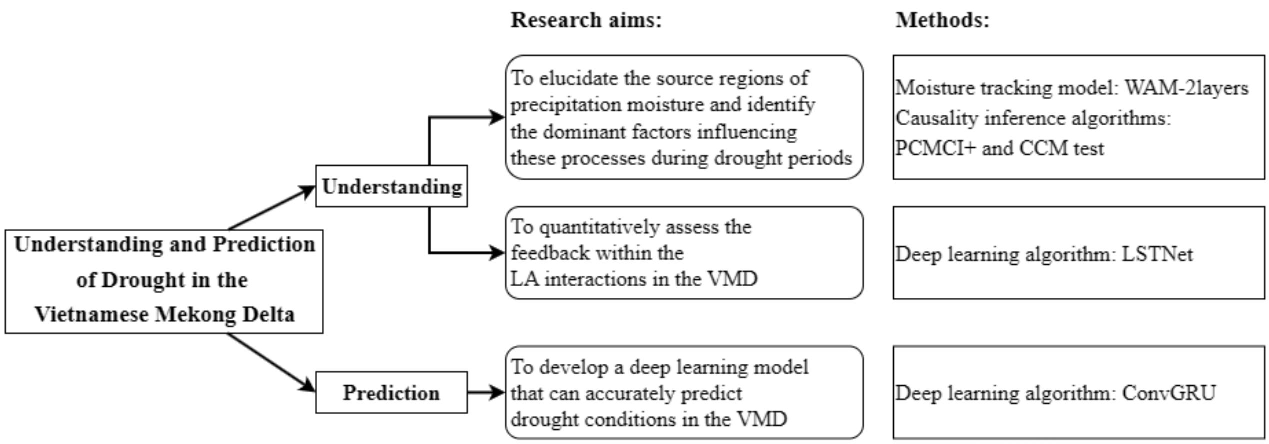

Figure 3.1. The general overview of the methods used to achieve the research aims and objectives in this study.

Chapter 4. Understanding precipitation moisture sources and their dominant factors during droughts in the Vietnamese Mekong Delta

Key Points:

- Precipitation moisture sources for the Vietnamese Mekong Delta primarily originate from the external area, accounting for over 60%.
- Based on causality algorithms, humidity and wind speed are the dominant factors of moisture transport in dry and wet seasons, respectively.
- The large-scale forcings and local atmospheric instability also have effects on moisture transport and recycling.

This chapter is a reformatted version of a manuscript published in the *Water Resources Research*, which is available at https://doi.org/10.1029/2023WR035920

Abstract

The Vietnamese Mekong Delta (VMD) is the most productive region in Vietnam in terms of agriculture and aquaculture. Unsurprisingly, droughts have been a prevalent concern for stakeholders across the VMD over the past decades. However, the VMD precipitation moisture sources and their dominant factors during drought conditions were not well understood. Using the ERA5 reanalysis data as inputs, the Water Accounting Model-2layers (WAM-2layers), a moisture tracking tool, was applied to identify the VMD precipitation moisture sources from 1980 to 2020. The modeling simulation indicates that the moisture sources transported from the upwind regions dominate the VMD precipitation by 60.4% to 93.3%, and the moisture source areas vary seasonally with different monsoon types. Results of the causal inference algorithms indicate that the humidity and wind speed in the upwind area are the dominant factors for driving moisture transport and determining the amount of VMD precipitation in dry and wet seasons, respectively. The local atmospheric conditions may also have a causal effect on moisture recycling. During the drought events in 2015-2016 and 2019-2020, the reduced moisture transport in the 2016 dry season was mainly caused by the anomalies in both humidity and wind speed, while the below average moisture sources in the 2020 dry season were dominated by humidity. In the 2019 wet season, an anomaly in wind speed led to a decrease in the tracked moisture. These findings are of great significance for understanding the moisture sources of precipitation and further improving drought prediction in the VMD.

4.1 Introduction

As one of the most destructive climate-related natural hazards, drought significantly impacts agriculture, ecology, and society (Wilhite and Glantz, 1985; Mishra and Singh, 2010; Vicca et al., 2016; Miralles et al., 2019). Over the past decades, the Mekong River Basin has experienced numerous severe drought events, notably in 1991-1994, 1998, 2005, 2010, 2015-2016, and 2019-2020, leading to serious social and economic consequences (Guo et al., 2017; Lee and Dang, 2018; Kang and Sridhar, 2021; Kang et al., 2021; Keovilignavong et al., 2021). Particularly, the droughts in 2015-2016 and 2019-2020 were devastating for the Vietnamese Mekong Delta (VMD), which is the most productive region in terms of agriculture and aquaculture in Vietnam. For example, the 2015-2016 drought caused a total of $360 million losses across all fields (Nguyen, 2017). In this tidal delta, seawater intrusion into inland areas during dry periods leads to widespread destruction of cropland, causing significant damage to the ecosystem and the livelihoods of local communities. Specifically, the 2015-2016 drought led to severe salinity intrusion, affecting over one million inhabitants due to the clean water shortage and damaging over 300,000 hectares of farmlands (United Nations Resident Coordinator in Viet Nam, 2016; Nguyen, 2017). During the 2019-2020 period, the extent of saltwater intrusion even surpassed the levels observed during the 2015-2016 drought (United Nations Resident Coordinator in Viet Nam, 2020).

After these two historic events, drought has emerged as a critical issue warranting more investigations in recent years in the VMD. For instance, Guo et al. (2017), Lee and Dang (2018), and Kang and Sridhar (2021) explored different types of historic droughts using various drought indices. Sam et al. (2019), Li Y. et al. (2021), Dong et al. (2022), and Kang et al. (2022) assessed the future drought characteristics in the Mekong River Basin, and the studies revealed that the frequency, duration, and severity of drought in this area are projected to increase in the near future. As for the drivers of drought events in this area, Lu et al. (2014) and Kantoush et al. (2017) believed the construction of dams can reduce the downstream discharge and lead to droughts. Cosslett and Cosslett (2018) pointed out that rainfall during EI Niño is generally below average with the wet season being shorter, resulting in record droughts in 2010 and 2015. Bastakoti et al. (2014) and Miyan (2015) attributed the occurrence of droughts to high temperature and low precipitation. Although the causes of drought are usually complex, the deficiency of precipitation is always the dominant factor leading to drought (Mishra and Singh, 2010). Therefore, the role of moisture recycling and transport (advection) in characterizing drought events (the deficiency of precipitation) has been an important research area (Roy et al., 2019).

Accordingly, moisture tracking models were developed rapidly in recent years, which are essential for comprehending drought dynamics. Notable developments include the Flexible Particle (FLEXPART, Sun and Wang, 2014), Hybrid- Single Particle Lagrangian Integrated Trajectory (HYSPLIT, Chu et al., 2017; Guan et al., 2022), Eulerian-based Water Accounting Model-2layers (WAM-2layers, Van der Ent et al., 2010; Van der Ent et al., 2014; Keys et al., 2014; Xiao and Cui, 2021), and Dynamical Recycling model (DRM, Dominguez et al., 2006; Cheng and Lu, 2020). Utilizing DRM, Herrera‐Estrada et al. (2019) investigated how diminished moisture transport influences drought propagation in North America. They found that dry soil moisture and reduced evapotranspiration in upwind land areas led to a decrease in moisture transport to the downwind areas, thereby exacerbating agricultural drought conditions in the North American subregions. Based on the outcomes of a Lagrangian back-trajectory approach, Holgate et al. (2020) found that the occurrence and intensification of drought in southeast Australia were usually dominated by the reduction of moisture from the ocean. Meanwhile, Benedict et al. (2021), Zhang (2020), and Guan et al. (2022) tracked the anomalies of moisture sources during the severe historic droughts in the Rhine basin, Southwest China, and Mid-to-Lower Reaches of the Yangtze River, respectively. As for the moisture tracking research in the VMD, Stojanovic et al. (2021) explored the moisture sources over the seven climate subregions of Vietnam based on the FLEXPART model. Their research revealed that precipitation moisture for the VMD is mainly from the China Seas in the dry season and the Bay of Bengal in the wet season. For this drought-prone delta, tracing and quantifying moisture sources of precipitation are essential for understanding drought generation mechanisms and forecasting the occurrence and severity of droughts. This is vital and valuable to strengthen drought preparedness and resilience in this area and support decision-makers in making optimal drought management and mitigation strategies.

Utilizing outputs from moisture tracking models and anomaly analyses of atmospheric fields (e.g., specific humidity, moisture flux), Benedict et al. (2021) and Guan et al. (2022) illustrated the connections between these variables and anomalies in moisture transport. Stojanovic et al. (2021) found a significant positive correlation between drought conditions and contribution moisture anomalies. In this study, for the first time, the causality inference algorithms were introduced to analyze the causal relationships among variables involved in moisture transport, specifically, to identify which factor drives the moisture transport process and dominates the amount of tracked moisture. Traditionally, correlation and regression are commonly used to characterize the dependence between variables in hydrometeorology research (e.g., Stojanovic et al., 2021). However, these methods fall short in addressing cause-and-effect relationships in the hydrometeorology system (Runge et al., 2019a; Ombadi et al., 2020). With the accumulation of a large number of in-situ and remote sensing hydrometeorological data, it is possible to infer the causal relations among variables through the causality inference algorithms, which have seen rapid development (Runge et al., 2019a). The Granger causality test (Granger, 1969), the earliest practical and empirical causal inference method based on linear vector autoregressive, stands out for its simplicity and wide application across various scientific domains (Shojaie and Fox, 2022). Its applicability has been extended to hydrometeorological research in recent years (Green et al., 2017; Tuttle and Salvucci, 2017). The GC test assumes that the variables are stochastic, however, to the extent that they are not entirely random. Therefore, the Convergent Cross Mapping (CCM, Sugihara et al., 2012) test was proposed to overcome this limitation and evaluate the causal relationships among variables in dynamical systems. It has been used to explore the causal effect of soil moisture on precipitation (Wang et al., 2018). Nevertheless, Ombadi et al. (2020) pointed out that CCM may yield wrong bidirectional results (actually unidirectional) when variables are strongly coupled. In addition, the graph-based Peter Clark (PC) algorithm (Spirtes and Glymour, 1991), named after its inventors Peter Spirtes and Clark Glymour, is also a popular causal inference method. The PC algorithm was applied to analyze the environmental drivers of evapotranspiration in the research of Ombadi et al. (2020). It has been further developed into the Peter and Clark Momentary Conditional Independence (PCMCI) and PCMCI+ algorithms by Runge et al. (2019b) and Runge (2020), which further conduct a Momentary Conditional Independence (MCI) test after the PC algorithm. The application of causality inference algorithms substantially clarifies understanding of the physical process in hydrometeorology and aids in deducing the dominant factors that drive the process at different stages. The causality network would be strong and clear evidence for the further prediction of drought in the VMD.

Because the WAM-2layers model is timesaving for long-term and large-scale moisture tracking, it was employed in this study to explore the role of moisture from upwind in the propagation of drought in the VMD. Then, the causality directions of the PCMCI+ algorithm and the causality strength of the CCM were integrated to comprehensively evaluate the causal relations of variables in the moisture transport process. To summarize, the primary objectives of this study are: 1) to identify and characterize the precipitationsheds for the VMD and explore the role of moisture from upwind in the drought propagation; 2) to identify the key factors that drive moisture transport and dominate the VMD precipitation; 3) to characterize two recent extreme drought events in the VMD based on the outputs of moisture tracking model.

4.2 Study Area, Data and Methods

4.2.1 Study Area

The VMD (Figure 4.1) consists of 13 provinces, located in the southernmost part (8.56°-11.03°N, 104.44°-106.82°E) of Vietnam. With approximately 4 million hectares of land area (Loc et al., 2021) and 18 million residents, the VMD contributed up to 18% of Vietnam’s GDP in 2018, mainly from the aquacultural and agricultural production (Tran et al., 2019). In this region, the monthly temperature varies from 26.1 ℃ to 29.0 ℃ and the averaged annual precipitation is approximately 2000 mm (calculated based on ERA5 data from 1980 to 2020, Figure 4.1(c)). The VMD is located in the tropical monsoon region, where the northeast monsoon prevails during the dry season and the southwest monsoon prevails in the wet season (Vu et al., 2018). The periods from December to April are defined as the dry seasons while the wet seasons are from June to October. May and November are the transitional months between the dry and wet seasons.

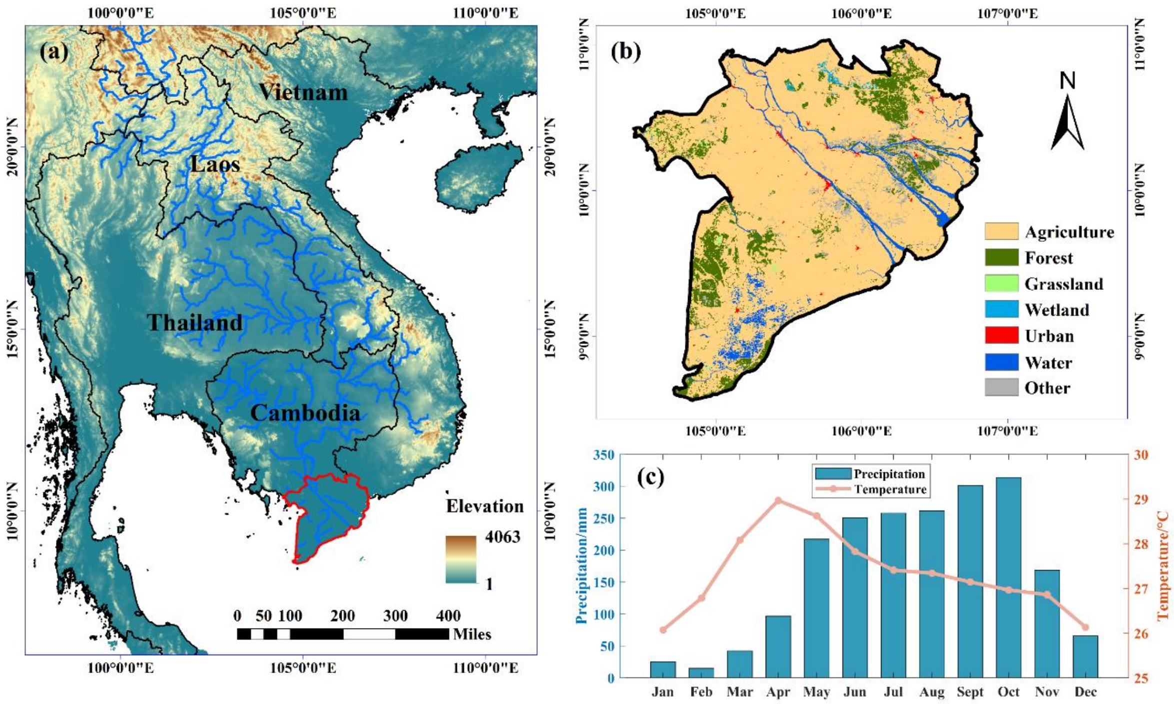

Figure 4.1. The (a) geographic location of Vietnamese Mekong Delta (VMD, within the red boundary), (b) land cover types derived from the European Space Agency Climate Change Initiative land cover datasets, and (c) precipitation and temperature climatology of VMD.

4.2.2 Data

The ERA5 reanalysis data were used in this study, including specific humidity, U- and V-components of wind at 17 pressure levels (i.e., 10, 100, 200, 300, 400, 500, 600, 700, 800, 825, 850, 875, 900, 925, 950, 975 and 1000 hPa), and total column water (TCW), precipitation, and evaporation at single level. The precipitation and evaporation were downloaded at an interval of 1 hour, while the remaining data were obtained at an interval of 6 hours. The ERA5 spatial resolution is 0.25°×0.25°. The data time span used here ranges from 1980 to 2020. In addition to hourly data, monthly vertical integral of divergence of moisture flux and vertical integral of water vapor flux derived from ERA5 were also used for the analyses (Hersbach et al., 2023a; 2023b). The ERA5 reanalysis data used in this study are freely available from the Copernicus Climate Change Service (C3S) Climate Data Store (https://cds.climate.copernicus.eu/cdsapp#!/home).

To examine the influence of large-scale climatic forcings on moisture transport processes and precipitation in the VMD, the Oceanic Niño Index (ONI) was employed to characterize the ENSO phenomenon. The ONI represents a three-month running mean of sea surface temperature (SST) anomalies within the Niño 3.4 region (120°W–170°W, 5°S–5°N). The dataset was obtained from the Climate Prediction Center (CPC). (https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php).

4.2.3 Causality inference methods

To investigate the causal relation among the variables in the moisture transport process, the following causality algorithms (i.e., PCMCI+ algorithm and CCM) were used in this study. The CCM test has the advantage of evaluating causality among variables in dynamical systems (e.g., weather systems), however, the CCM's potential for erroneous bidirectional inferences in the cases of strong variable coupling has been pointed out (Ombadi et al., 2020). In this study, therefore, the causality strength from the CCM test and the causality structure from the PCMCI+ algorithm were coupled to utilize the advantages of different causal inference algorithms. The code of the PCMCI+ algorithm was shared by Runge (2022) on GitHub (https://github.com/jakobrunge/tigramite), and the code of the CCM algorithm was shared by Javier (2021) on GitHub (https://github.com/PrinceJavier/causal_ccm).

4.2.3.1 Peter and Clark momentary conditional independence plus algorithm

The PCMCI algorithm is a causal discovery algorithm based on the causal graph modelling framework proposed by Runge et al. (2019a). The method couples the PC algorithm and MCI test to identify the presence of causal relationships by linking causal discovery to a causal inference framework, thus enabling more reliable causal network estimation, and producing more accurate causal effects (Runge, 2018; Runge et al., 2019b; Krich et al., 2020; Nowack et al., 2020). The PCMCI+ algorithm (Runge, 2020) was developed from PC and PCMCI algorithms, which starts from the PC algorithm in the first step and conducts the MCI test to address false positives. To overcome the disadvantage that the PCMCI cannot discover the contemporaneous causal links, the PCMCI+ was proposed. The brief process of the PCMCI+ algorithm is as follows:

In the first phase, the PC algorithm (three iterations in Figure 4.2(a), including p = 0, p = 1 and p = 2) is employed to establish and identify pseudo-links between variables and their respective sets of parent nodes. The set of parent nodes for each variable is initialized as follows:

$$\widehat{B_t^-}(X_t^j) = (X_{t-1}, X_{t-2}, \dots, X_{t-\tau_{max}}) \tag{4.1}$$

where $\widehat{B_t^-}(X_t^j)$ represents the set of parent nodes (lagged adjacency) for the variable X^j, and τ is the time lag. As illustrated in Figure 4.2(a), lagged links are oriented forward temporally (causes precede effects), while contemporaneous links remain undirected during the first phase. In the first iteration (p = 0), variables that fail the unconditional dependency test (e.g., uncorrelated) are removed from $\widehat{B_t^-}(X_t^j)$ (lightest shade of blue in Figure 4.2(a)). In the second iteration (p = 1), variables, which become independent conditional on the driver in $\widehat{B_t^-}(X_t^j)$, with largest dependency in the previous iteration are removed. In the third iteration (p = 2), variables are removed if they are independent conditionally on the two strongest drivers, and so forth until no further conditions can be tested in $\widehat{B_t^-}(X_t^j)$. Consequently, the PC algorithm adaptively converges to typically only a few relevant conditions (dark blue) that include the causal parents with high probability, along with potential false positives.

In the second phase, the lagged links (grey arrows) and contemporaneous links (undirected grey lines) will be tested with the MCI test, the pseudo-links (false positive) will be removed, and the true causal relationships will be oriented and obtained. All links (lagged and contemporaneous) from the first phase are tested with the MCI test:

$$X_{t-\tau}^i \perp\!\!\!\perp X_t^j \,|\, S, \widehat{B_t^-}(X_t^j) \setminus \{X_{t-\tau}^i\}, \widehat{B_t^-}(X_{t-\tau}^i) \tag{4.2}$$

where contemporaneous conditions $S \subseteq A_t(X_t^j)$, and $A_t(X_t^j)$ means contemporaneous adjacencies, here, S blocks contemporaneous paths, $\widehat{B_t^-}(X_t^j)$ and $\widehat{B_t^-}(X_{t-\tau}^i)$ block lagged paths.

The detailed description of PC, PCMCI and PCMCI+ algorithms can be found in Kalisch and Buhlmann (2007), Runge et al. (2019a, b), Ombadi et al. (2020) and Runge (2020).

4.2.3.2 Convergent cross mapping test

Unlike other methods that assume variables are stochastic, the CCM test assumes that the interaction takes place in a potential dynamic system and attempts to reveal the causal relationship based on Takens' theorem (Takens, 1981). It is usually considered as a supplement to the more statistical methods (Sugihara et al., 2012; Runge et al., 2019a). In dynamical systems theory, variables can be causally linked if they share a common dynamic system (Deyle and Sugihara, 2011), which means each variable can be used to identify the state of another variable. For example, points nearby on M_x will correspond temporally to points that are nearby on M_y if variables X and Y are dynamically coupled (Figure 4.2(b)). Therefore, to evaluate if variable Y has a causal effect on X, first, the shadow manifold M_x is used to identify the nearest neighbors of the point M_{xt} and their Euclidian distance from the point M_{xt}, which are used to estimate the contemporary state of M_{yt}. Variable Y has a causal effect on X if the estimated and observed states are significantly correlated. The central idea of CCM is that the cause leaves an information signature in the time series of the effect. The detailed CCM processes can be found in Sugihara et al. (2012) and Ombadi et al. (2020).

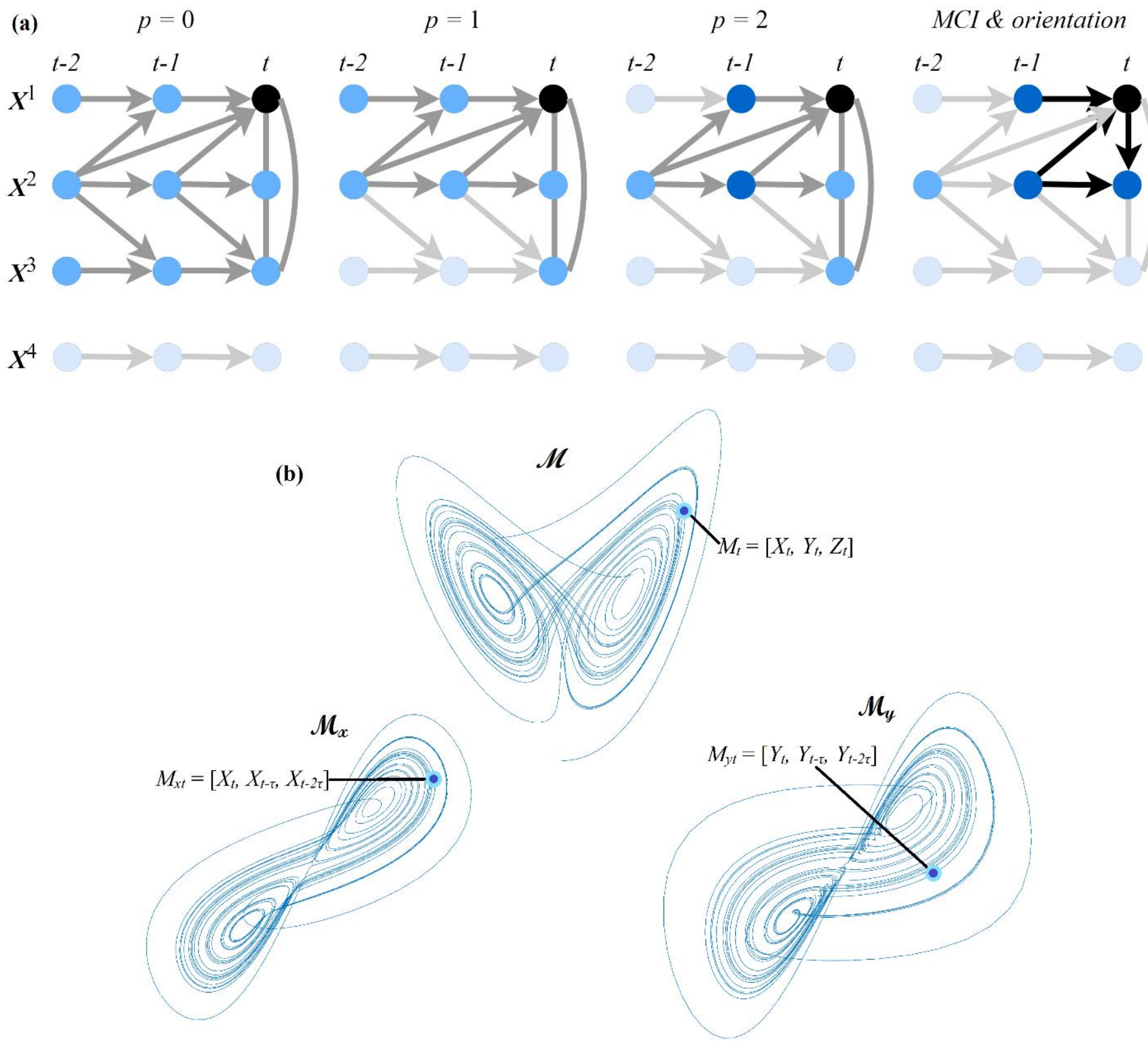

Figure 4.2. (a) The iteration phases in the PCMCI+ algorithm for identifying causal connections (the figure was reproduced based on Runge et al. (2019b)); (b) The original

manifold of the Lorenz system, M, and two shadow manifolds M_x and M_y, which are constructed from the variables (i.e., X, Y) and their lagged values (lag = τ). The manifolds are based on the Lorenz system with the equations: ${dx}/{dt} = 10(y - x), {dy}/{dt} = 28x - xz - y, {dz}/{dt} = xy - \frac{8}{3}z$ and initial values of x, y and z are 0.1 (the figure was reproduced based on Ombadi et al. (2020)).

4.2.4 Water accounting model with 2 layers

The Eulerian-based WAM-2layers model, developed by Van der Ent et al. (2013, 2014), is designed to track the tagged moisture (contribution evaporation) of the precipitation in the target area on both regional and global scales. It is an updated version of the Water Accounting Model (Van der Ent et al., 2010), and the two layers (bottom and top) in the WAM-2layers model were set to resolve the problem of wind shear in the upper air. The WAM-2layers offers rapid calculation capabilities for large-scale and long-term atmospheric moisture tracking (Van der Ent, 2014; Zhang, 2020; Li Y. et al., 2022). The tracking principle is based on the atmospheric water balance (Van der Ent et al., 2014):

$$\frac{\partial S_k}{\partial t} = \frac{\partial (S_k u)}{\partial x} + \frac{\partial (S_k v)}{\partial y} + E_k - P_k + \varepsilon_k \pm F_v \tag{4.3}$$

where S_k indicates the atmospheric moisture storage in layer k (the top or bottom layer), E and P indicate the evaporation and precipitation, u and v indicate the wind speed in the zonal- and meridional-direction, respectively, ε indicates the residual, and F_v is the vertical moisture transport; $\frac{\partial (S_k u)}{\partial x}$ and $\frac{\partial (S_k v)}{\partial y}$ represent the horizontal moisture transport, which can be calculated with moisture flux F_k. F_k can be described as follows (Van der Ent et al., 2014):

$$F_k = \frac{L}{g\rho_w} \int_{p_{top}}^{p_{bottom}} q u_h dp \tag{4.4}$$

where L is the length of the grid cell perpendicular to the direction of the moisture flux, g is the gravitational acceleration, ρ_w the density of liquid water (1000 kgm^{-3}), p stands for pressure, q stands for specific humidity, and u_h is the horizontal component in either zonal or meridional direction.

In the WAM-2layers, the column air was divided into the bottom and top layers by P_{divide}, which is calculated by Van der Ent et al. (2014):

$$p_{divide} = 7438.803 + 0.728786 \times p_{surface} \quad (4.5)$$

where $P_{surface}$ indicates surface pressure. In this study, the pressure-level data were derived from ERA5. The division of top and bottom layers can be seen from Xiao and Cui (2021).

The main output of the WAM-2layers model is the distribution of contribution evaporation (i.e., tagged moisture) that travels through the atmosphere and contributes to the precipitation in the target (sink) region. Based on this, the concept of precipitationshed was proposed to illustrate how the upwind evaporation source areas contribute moisture for precipitation to the downwind sink regions (Keys et al., 2012; Keys et al., 2014). According to the WAM-2layers model, the precipitationshed should cover the entire globe. However, since the contribution evaporation in most grid cells is very small, a threshold should be set to determine the proper boundary of the precipitationshed. Keys et al. (2012) defined the threshold as 70% of growing season precipitation was contributed by the moisture from the precipitationshed. Xiao and Cui (2021) defined that a grid cell is considered to be in the precipitationshed if the contribution evaporation of the grid cell exceeds 0.02% of its total evaporation. Given that the study area is only for VMD, it is not necessary to take the whole globe as the evaporation source region in the model. Therefore, the domain of the source region is defined at latitude from 60°N to 30°S, and longitude from 60°E to 150°E. In this study, the definition of precipitationshed was adopted from Zhang (2020). Precipitationshed consists of grid cells with contribution evaporation higher than the threshold (e.g., 0.28 mm in the dry season in Figure 4.3), which sum up to 90% of the total contribution evaporation from the source region. Here, monthly and seasonal precipitationsheds for the VMD were obtained based on this definition. Figure 4.3 represents the seasonal precipitationsheds in both dry and wet seasons. The precipitationsheds were divided into local (inside of the VMD) and external (outside of the VMD) parts for further analyses.

The code of WAM-2layers model was shared by xmingzh (2021) at Zenodo (http://doi.org/10.5281/zenodo.4796962).

4.3 Results

4.3.1 Features of precipitationshed for the VMD

The monthly and seasonal precipitationsheds for the VMD were identified based on outputs from the WAM-2layers model. Table 4.1 presents the climatological monthly contribution evaporation for the VMD precipitation from the local and external areas, along with the recycling ratio of local total evaporation. It is evident that the contribution evaporation from the external area significantly exceeds that from the local area each month. The monthly local recycling ratio ranges from about 1% to 24%, which in the wet season is higher than in the dry season. Figure 4.3 shows the seasonal precipitationsheds for the dry and wet seasons. The amount of contribution evaporation within the precipitationshed ranges from 0.28 mm to 19.47 mm during the dry season and from 0.83 mm to 106.17 mm in the wet season. Located in the monsoon area, the VMD experiences the northeast monsoon in dry seasons and the southwest monsoon during wet seasons. It is clear from the seasonal precipitationsheds that most of the contribution evaporation in dry seasons originates from the northeast areas (e.g., the South China Sea), while in wet seasons, the southwest areas (e.g., the Bay of Bengal) are the primary moisture supplier for the precipitation. Due to the effect of precipitation sinking during the moisture transport, the closer to the target region (i.e., the VMD, within the red boundary in Figure 4.3), and the more evaporation is contributed from the sources. To evaluate the performance of the WAM-2layers model, the amount of contribution evaporation from the identified precipitationsheds was compared with the precipitation data from ERA5. Figure 4.4(a) demonstrates a strong correlation between the contribution evaporation and the ERA5 precipitation data. The model tracked approximately 70% of precipitation moisture from the precipitationsheds, which aligns with the precipitationshed definition by Keys et al. (2012) as described in Section 4.2.4. From 1980 to 2020, contribution evaporation from the precipitationsheds contributed 62.8%-97.0% of the total contribution evaporation, with the local area accounting for only 1.2%-27.1%, while the external areas contributed between 60.4% and 93.3% (Figure 4.4(b)). Moisture from source regions outside the precipitationsheds (represented by the residual component in Figure 4.4(b)) accounted for 3.0%-37.2% of the total contribution evaporation. Due to the effect of monsoon, the moisture from the external area is the primary source of the precipitation in the VMD.

Figure 4.4(c) reveals that the anomalies in contribution evaporation correspond closely to the historical drought events in the VMD, e.g., drought events in 2002, 2005, 2015, and 2019 recorded by the Emergency Events Database (EM-DAT), as well as the 1990-1994 and 1998 drought events identified by previous studies (Guo et al., 2017; Lee and Dang, 2018; Lee and Dang, 2019; Le et al., 2020). The deficiency of moisture from the external area accounts for the main part of the deficits in drought years, which is consistent with the contribution rate in Figure 4.4(b).

Table 4.1. The monthly climatology contribution evaporation from the local and external areas and the recycling ratio of local total evaporation.

Month		Jan	Feb	Mar	Apr	May	Jun
Contribution Evaporation/mm (%)	**External**	16.25	14.09	33.12	84.17	156.75	148.83
	Local	1.48 (1.48)	0.77 (0.87)	2.95 (3.01)	7.00 (6.46)	18.99 (15.17)	13.89 (11.21)
Month		**Jul**	**Aug**	**Sept**	**Oct**	**Nov**	**Dec**
Contribution Evaporation/mm (%)	**External**	177.29	168.58	219.81	156.56	80.66	35.77
	Local	10.05 (7.73)	9.17 (6.98)	20.91 (17.39)	28.16 (23.97)	7.67 (6.61)	2.15 (1.96)

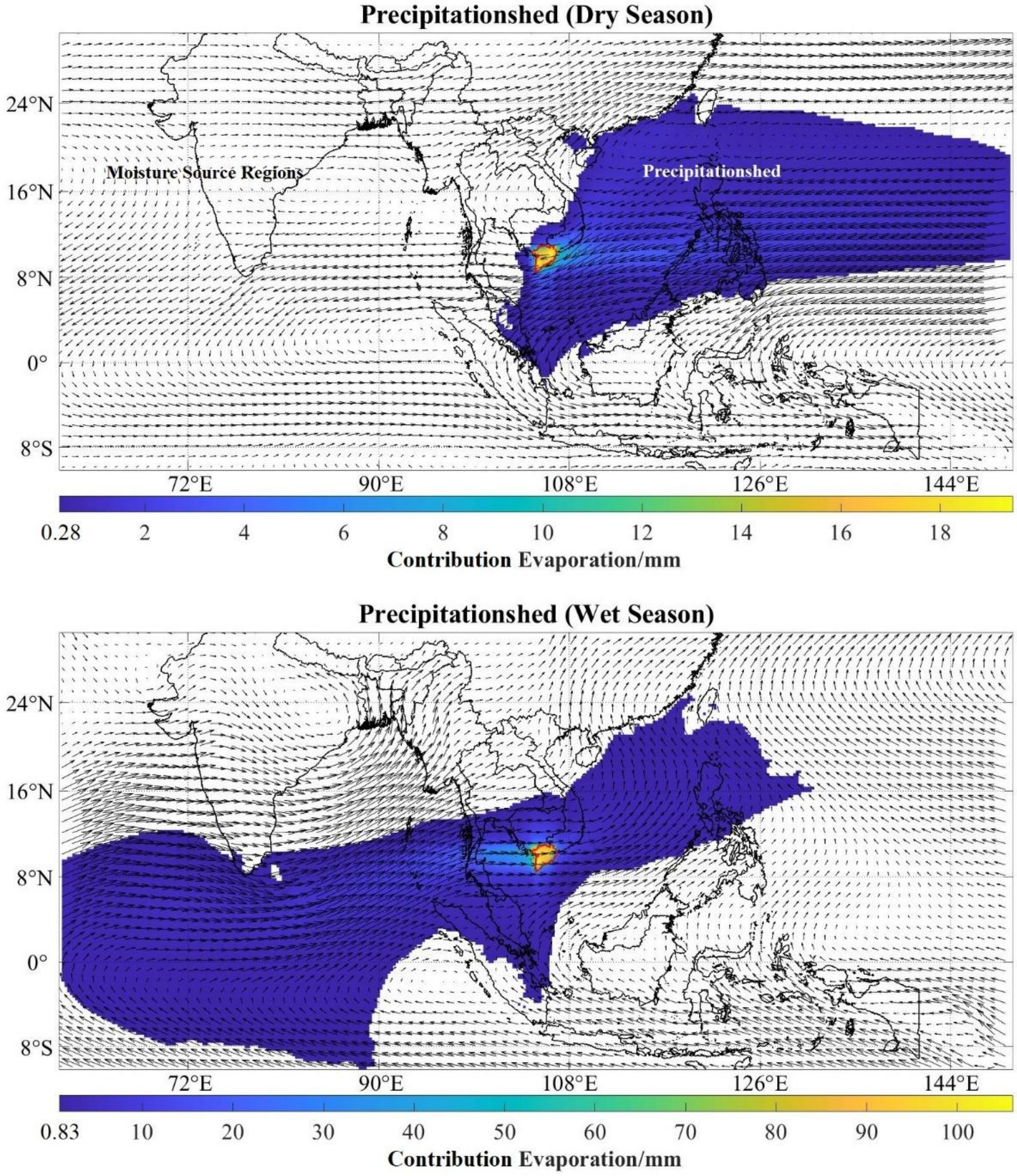

Figure 4.3. Dry and wet seasons precipitationshed in the VMD.

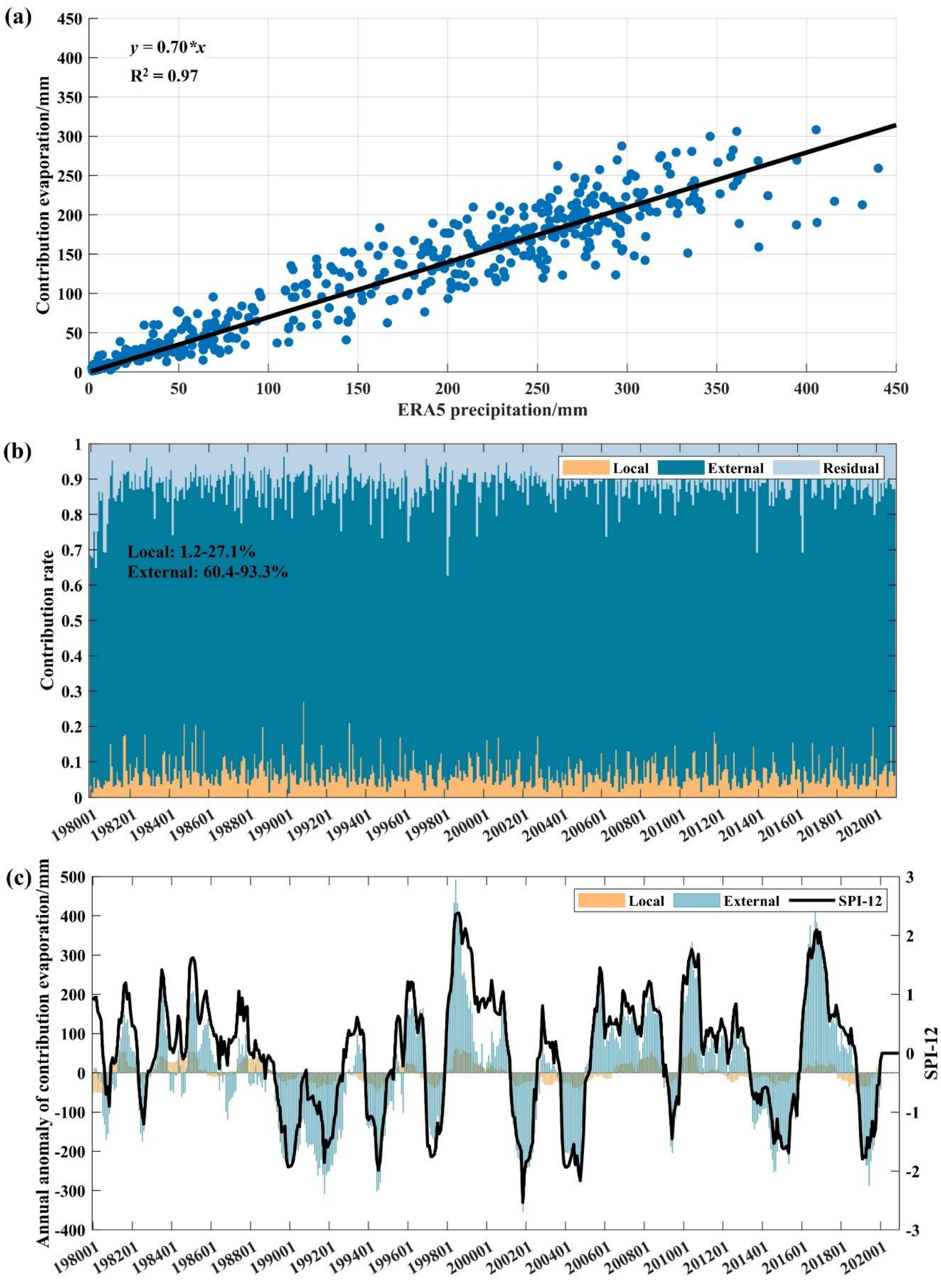

Figure 4.4. (a) Monthly contribution evaporation from WAM-2layers versus the monthly ERA5 precipitation; (b) Contribution rate of local and external contribution evaporation relative to all the tracked evaporation from the source region; (c) Annual anomaly of local and external contribution evaporation and time series of SPI-12 (12-month Standardized Precipitation Index).

4.3.2 Causal relation among the variables during the moisture transport

Correlation analysis is usually insufficient for deducing causal relationships among variables in the Earth system (Runge et al., 2019a; Ombadi et al., 2020). Therefore, the PCMCI+ algorithm was utilized to establish the direction of causal links, while the CCM test was used to determine the strength of these causal relationships. As shown in Figure 4.5(a), the causal relations among the variables during the moisture transport process are depicted based on these two algorithms. During moisture transport, humidity and wind speed are the two important factors that affect moisture flux. Therefore, in the causality analysis, TCW_local and TCW_external represented the vertical integral humidity of local and external areas, respectively. The zonal wind is prevalent in the upwind area. Therefore, U-wind is selected to illustrate the effects of wind on moisture transport. As the contribution evaporation from the external area accounts for most of the VMD precipitation, CE_external was selected to bridge external atmospheric conditions and VMD precipitation. The causality network varies between dry and wet seasons. In dry seasons, TCW_external influences the amount of contribution evaporation from the external area, subsequently impacting the amount of precipitation in the VMD. In wet seasons, the primary variable that affects CE_external is wind speed. Due to drier air conditions in the upwind area during dry seasons, the moisture flux is predominantly constrained by humidity. Therefore, in the causality network in the dry season, TCW is the critical factor that constrained the amount of CE_external. On the contrary, with sufficient water vapor (TCW) in the upwind area during wet seasons, the anomaly in wind speed dominates the amount of water vapor flux. Similarly, in Figure 4.6, the coefficient of determination (R^2) and Spearman's rank correlation coefficient (R_S) between TCW_external and CE_external is higher in the dry season (0.65 and 0.79) compared to the wet season (0.24 and 0.52), respectively.

The strength of the causal link between CE_external and Precipitation is higher in the dry season than in the wet season (0.87 vs 0.60), suggesting that contribution evaporation from the external area is more crucial to the VMD precipitation in dry seasons than in wet seasons. The time series of the SPI-12 and annual anomaly of contribution evaporation from the external area correlate very well (Figure 4.4), which supports the result of the causal link between CE_external and Precipitation. The causal links between TCW_local and Precipitation in dry and wet seasons imply the effect of local atmospheric conditions on the amount of precipitation. The strength of causal links between these two variables was stronger in the dry season (0.71) compared to the wet season (0.40).

Figure 4.5(b) presents the spatial distribution of causality strength between TCW, U-wind and contribution evaporation based on the CCM test in dry and wet seasons, and it was weighted by the contribution evaporation. The results here are consistent with Figure 4.5(a) that contribution evaporation is more sensitive to humidity (wind speed) in dry (wet) seasons.

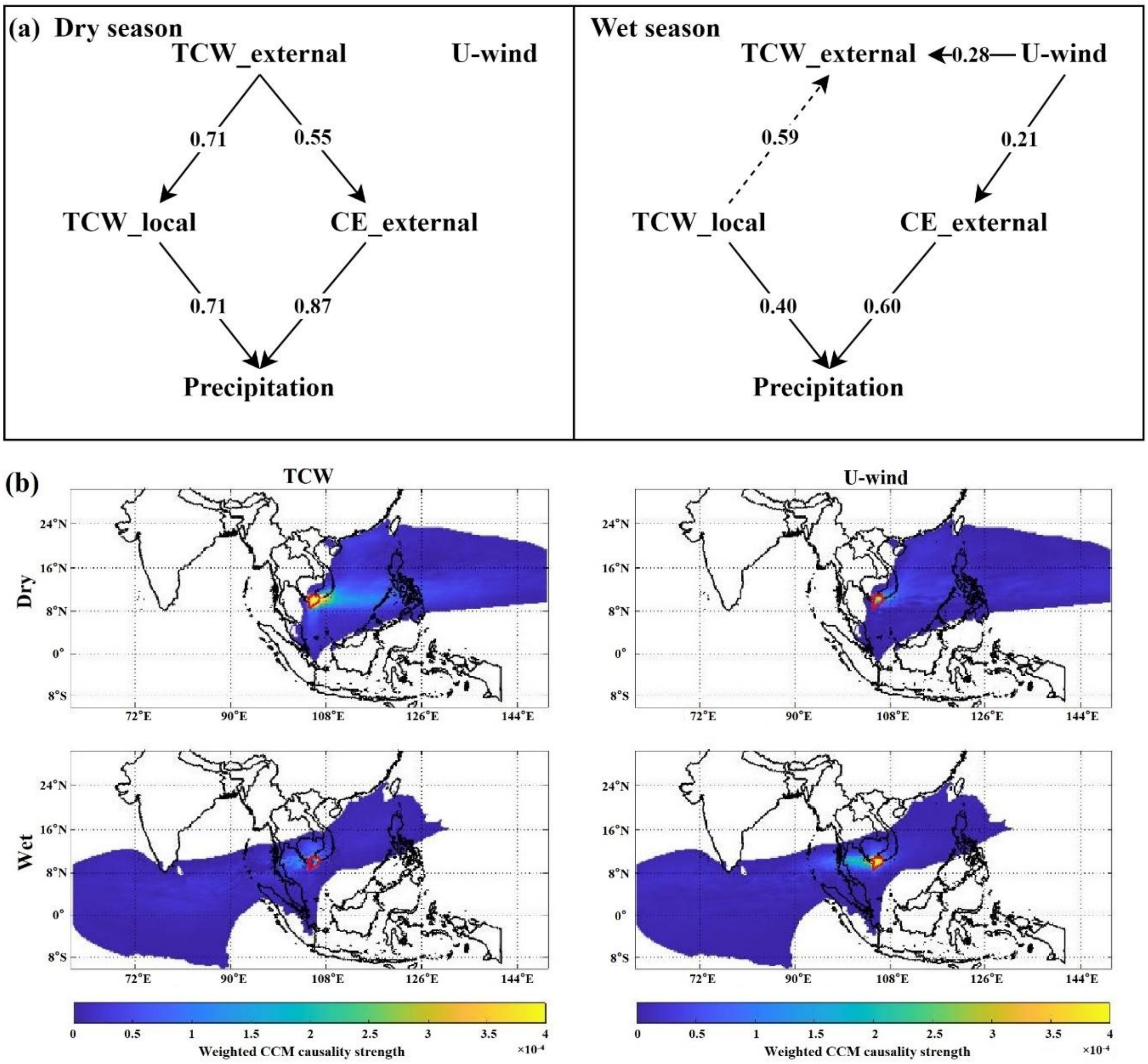

Figure 4.5. (a) Causal relations among the variables of moisture transportation. *TCW_external*: total column water in the external area, which is weighted averaged according to the contribution evaporation in each grid cell; *TCW_local*: total column water in the local area; *U-wind*: weighted averaged vertical integral U-component of wind in external area according to the contribution evaporation; *CE_external*: total amount of contribution evaporation from external area; *Precipitation*: the total amount of precipitation in the VMD. The dotted line indicates causal links falsely identified by the PCMCI+ algorithm. The figures of edges represent the strength of the causal links evaluated by the CCM; (b) The sensitivity of contribution evaporation to TCW (left) and U-wind (right) based on the CCM in dry and wet seasons respectively.

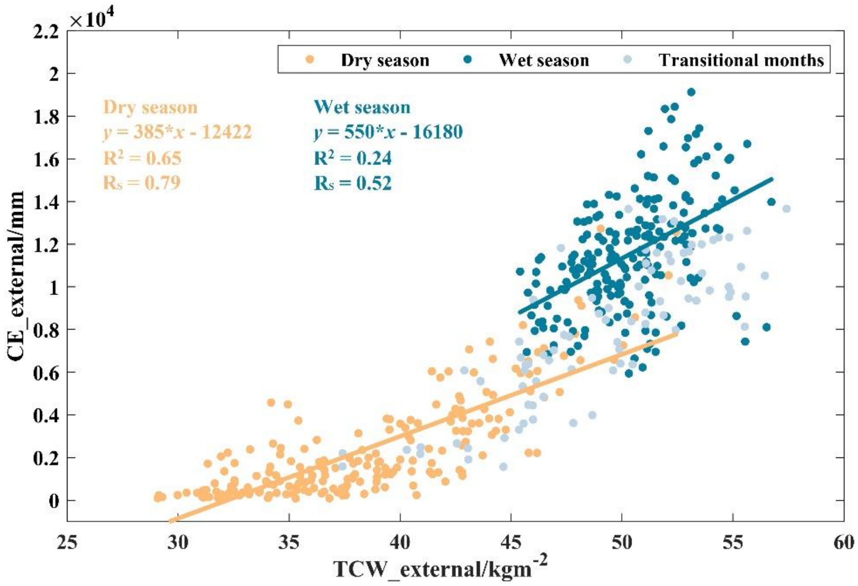

Figure 4.6. The correlation between TCW_external and CE_external in dry and wet season.

4.3.3 Moisture transport anomalies of two extreme drought events

Droughts in 2015-2016 and 2019-2020 are the two most severe VMD drought events in recent years (Guo et al., 2017; Frappart et al., 2018; Loc et al., 2021). Figure 4.7 presents the moisture transport anomalies of these two drought events.

In the 2015-2016 drought, the negative anomaly in contribution evaporation from the external area began in December 2014, and the negative anomaly of the local area started one month later. In the 2015 dry season (December 2014 to April 2015), local recycled evaporation decreased by 5.74 mm, accounting for 40.02% of the local climatology contribution evaporation. Meanwhile, the contribution evaporation from the external area decreased by 79.61 mm (43.41% of the external climatology contribution evaporation for the same period). From May to August 2015, the contribution evaporation continued to show negative anomalies, which decreased by 22.64 mm (43.46%) in the local area and 131.5 mm (20.19%) in the external area. Afterwards, the drought situation relieved slightly in the late wet season of 2015 (September 2015 to November 2015), with an increase in contribution evaporation of 40.44 mm (10.75%) in the upwind area in September and October 2015, followed by a decrease of 12.85 mm (15.94%) in November 2015. Unlike the external area, the positive anomaly in the VMD emerged one month later, with a decrease of 0.61 mm in September and an increase of 9.16 mm in October and November. The subsequent 2016 dry season (December 2015 to April 2016) was the driest period of this drought event. Only less than 20% of the climatology local contribution evaporation was recycled (decreased by 11.65 mm), in the meanwhile, the contribution evaporation from the upwind area decreased by 127.26 mm (69.49%). The negative anomaly in external contribution evaporation ended in April 2016, while the local recycled evaporation continued to decrease in May by 1.85 mm (9.74%). During this drought event, the variation in local contribution evaporation was one month delayed compared with the external contribution evaporation.

In the initial stage of the 2019-2020 drought (January to May 2019), the anomaly in local recycled evaporation also lagged one month behind the external anomaly. Compared with the climatological average, local contribution evaporation increased slightly in January 2019, then decreased by 11.96 mm (40.25% of the local climatology contribution evaporation) from February to May. For the external area, contribution evaporation reduced by 41.98 mm (28.44%) from January to April and increased by 18.87 mm (12.04%) in May. Subsequently, the local recycling ratio increased in June 2019 (2.24 mm, 16.09%). In the wet season, the contribution evaporation mainly exhibited a negative anomaly, except for a local positive anomaly in June. The contribution evaporation decreased by 19.05 mm (23.18%) in the VMD and by 101.04 mm (11.6%) in the upwind area. Unlike the alleviation of dry conditions in the late wet season of 2015, there was a negative anomaly in the 2019 wet season. The subsequent 2020 dry season (December 2019 to April 2020) was also the driest period of this drought. Only around half of climatological local contribution evaporation was recycled (decreased by 7.65 mm, 53.32%), in the meanwhile, the contribution evaporation from the external area decreased by 100.89 mm (55.01%). The 2019-2020 drought ended in May 2020 with a 6.96 mm (36.67%) and 33.45 mm (21.37%) decrease in contribution evaporation in the local and external areas, respectively.

The accumulated contribution evaporation anomalies of 2015-2016 and 2019-2020 droughts show similar spatial distributions (Figure 4.7(c)). Due to precipitation sinking during the moisture transport, the closer to the local region, the more evaporation contributed from the sources (Section 4.3.1) and the higher contribution evaporation anomaly. The positive contribution evaporation anomaly in the Bay of Bengal and Indian Ocean in the 2015-2016 drought corresponds to the relief of dry conditions in the late wet season of 2015. The contribution evaporation showed a negative anomaly in most of the wet precipitationsheds in 2019-2020. In most of the dry precipitationsheds (northeast regions), the anomaly of the 2015-2016 drought was lower than the 2019-2020 drought, which is consistent with the more severe drought conditions in the dry seasons of 2015-2016 than 2019-2020 (Figures 4.7(a) and (b)).

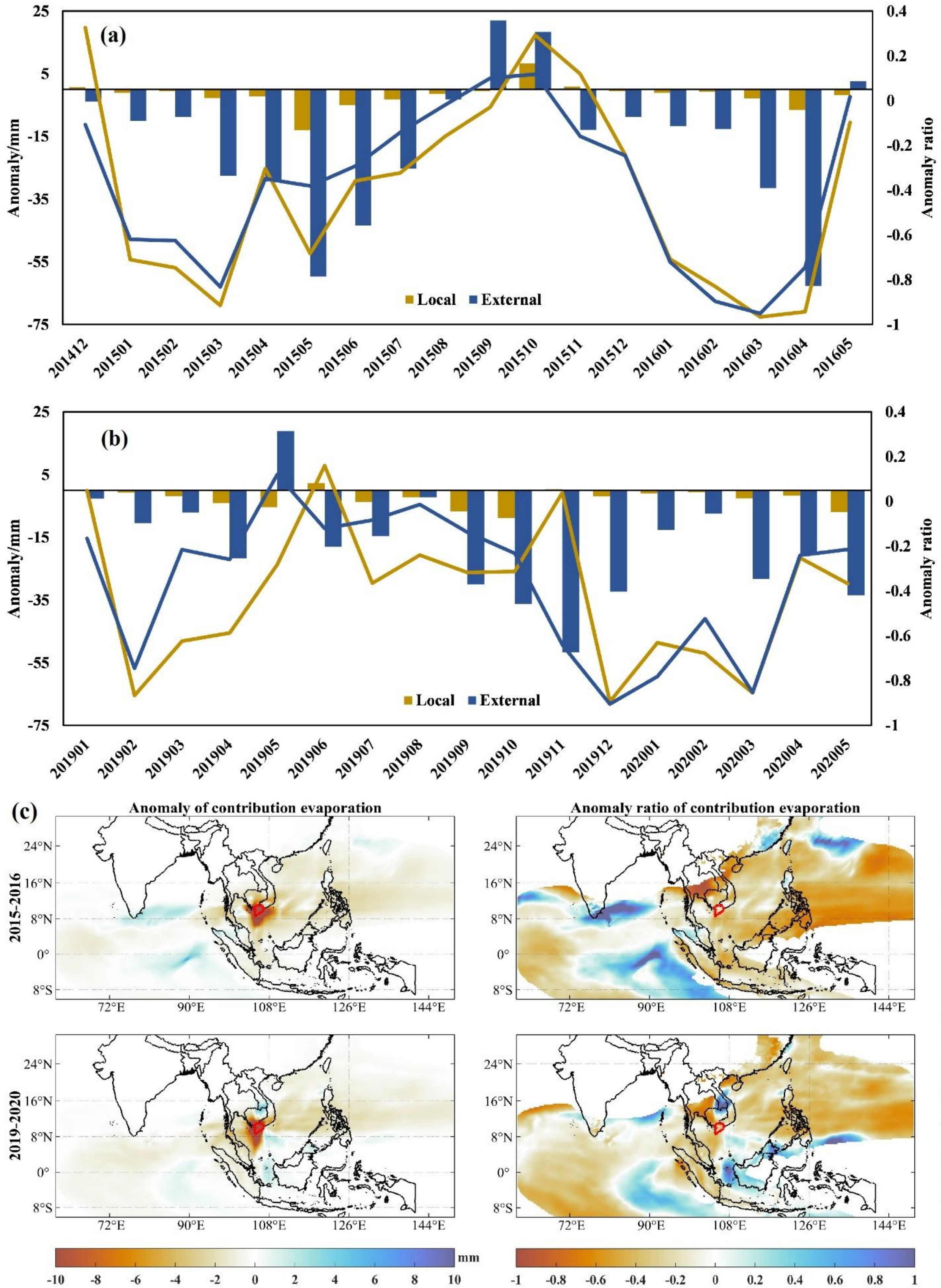

Figure 4.7. The variation of contribution evaporation anomaly in the local and external area (bar) and the anomaly ratio relative to the climatology (line) in (a) the 2015-2016 drought and

(b) the 2019-2020 drought; (c) Spatial distribution of accumulated contribution evaporation anomalies and their ratio relative to the climatology of 2015-2016 and 2019-2020 droughts.

Furthermore, anomalies in TCW and U-wind were calculated to identify factors that may affect moisture transport during these two significant drought events (Figures 4.8 and 4.9). Given that positive and negative U-wind values indicate eastward and westward directions, the wind speed is first normalized to the 0-1 range before calculating the anomalies. Positive and negative U-wind anomalies correspond to eastward and westward deviations, respectively. A positive eastward anomaly of U-wind in dry seasons means a weakened northeast monsoon, which could reduce the amount of tracked water vapor. For instance, during the 2016 dry season, under the influence of the positive eastward anomaly in U-wind and the negative anomaly in TCW, the contribution evaporation decreased by around 80%. In the 2020 dry season, because the U-wind showed a westward anomaly, the negative anomaly of contribution evaporation was primarily due to the drier condition in the upwind area. In the wet seasons, whether in the wet season of 2015 or 2019, the anomaly of TCW mainly ranges from -0.05 to 0.05 in most parts of the precipitationshed. The U-wind anomaly was slightly positive eastward in the 2015 wet season, which is consistent with the result that contribution evaporation in the 2015 wet season was slightly above the climatological average. On the contrary, the U-wind in the 2019 wet season showed a strong negative westward anomaly, aligning with the reduction of contribution evaporation.

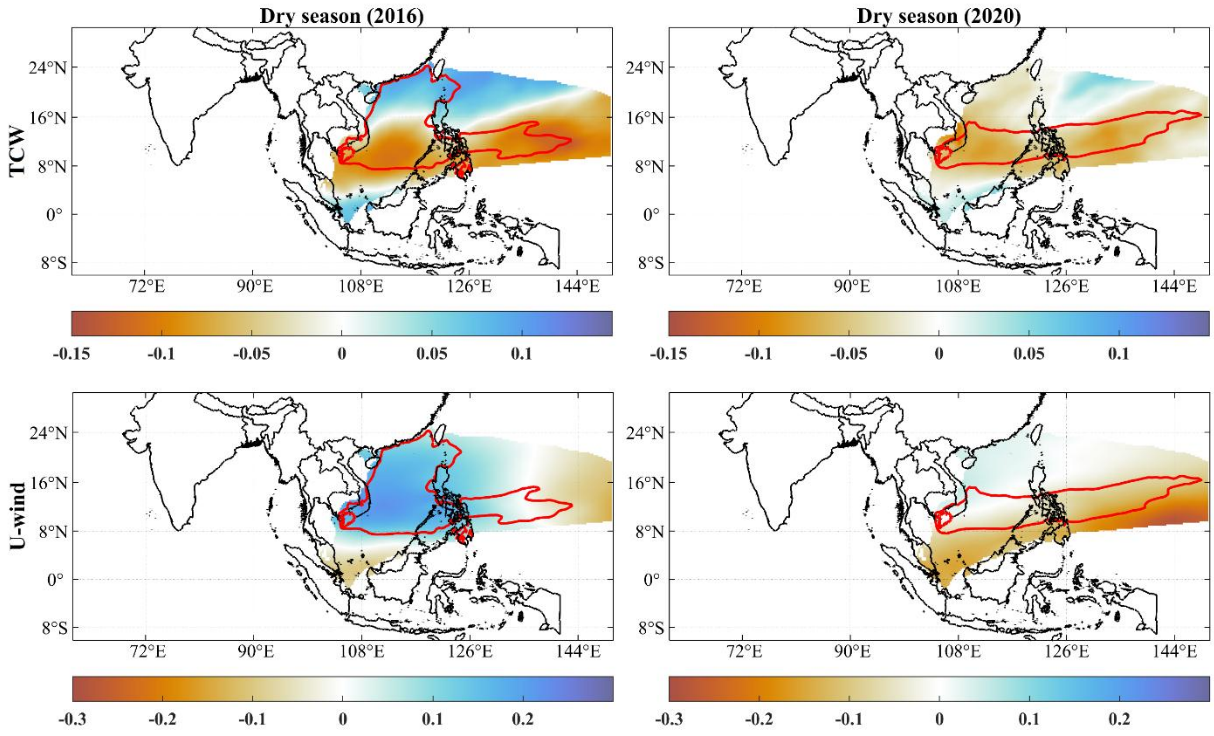

Figure 4.8. Anomalies of TCW and U-wind during the 2016 and 2020 dry seasons relative to the corresponding climatology. The red line indicates the core precipitationshed, which accounts for 70% contribution evaporation of the whole precipitationshed.

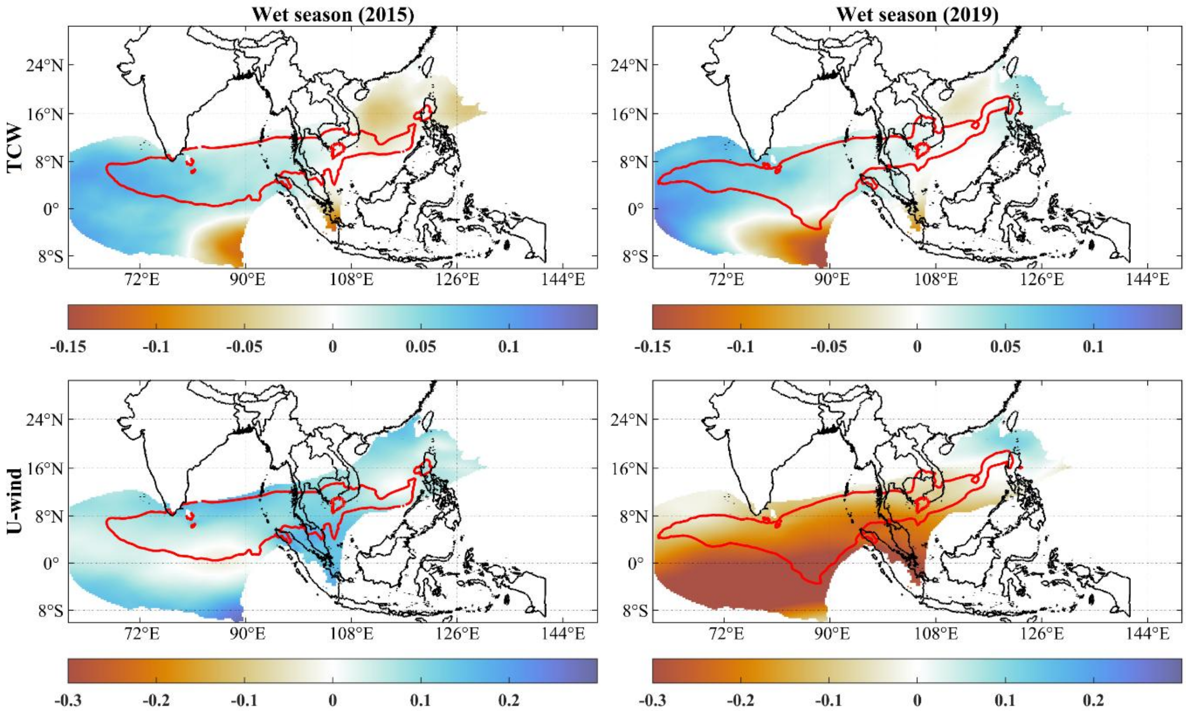

Figure 4.9. Anomalies of TCW and U-wind during the 2015 and 2019 wet seasons relative to the corresponding climatology. The red line indicates the core precipitationshed, which accounts for 70% contribution evaporation of the whole precipitationshed.

4.3.4 Effects of large-scale forcings and local atmospheric conditions on the anomalies of contribution evaporation

Regional moisture transport can be affected by large-scale forcings such as El Niño and La Niña. Stojanovic et al. (2021) found the associations between these phenomena and dry conditions in Vietnam differ across dry and wet seasons, as well as various subregions of Vietnam. ENSO represents one of the Earth's most significant modes of interannual climate variability, influencing global weather patterns through atmospheric teleconnections (Yeh et al., 2018, Xu et al., 2025). These teleconnections transmit ENSO's effects to distant regions, shaping the frequency and intensity of extreme weather events, such as droughts and floods, across various parts of the world (Vicente-Serrano et al., 2011). Given its widespread influence, this section examines the extent to which ENSO phenomena modulate the moisture transport mechanisms that contribute to precipitation in the VMD region, providing insights into its role in regional hydroclimate variability. ONI was utilized to represent the variation of

SST in the Niño 3.4 region (120°W-170°W and 5°S-5°N). We identified the dry and wet seasons during which El Niño and La Niña phenomena occur, as well as the normal dry and wet seasons. Figure 4.10(a) shows the ratio between the anomalies in the contribution evaporation at different phases (i.e., El Niño and La Niña) and the contribution evaporation in the normal dry or wet seasons. It reveals that the amount of contribution evaporation exhibits a negative anomaly in the El Niño phase but a strong positive anomaly in the La Niña phase across the precipitationshed. Generally, the effect of ENSO phenomena on the amount of contribution evaporation in the wet season is not consistent with the dry season across the precipitationshed.

Given that TCW and U-wind play important roles in the moisture transportation during dry and wet seasons, respectively, the difference in the spatial pattern of TCW in the dry season and U-wind in the wet season under different ENSO phases was investigated (Figures 4.10(b) and (c)). The atmosphere over the precipitationshed would be drier (wetter) in the El Niño (La Niña) phase, which corresponds to the negative (positive) anomaly in the contribution evaporation under different ENSO phases. This result is consistent with the causality network as described in Section 4.3.2. On the other hand, the U-wind shows a negative anomaly across the precipitationshed during the La Niña phase. During the El Niño phase, the U-wind near the equator is below average while it shows a slightly positive anomaly in the northern precipitationshed. Comparing the results in the wet season in Figures 4.10(a) and (c), the positive or negative anomalies of U-wind in different areas may have different effects on the amount of contribution evaporation. For example, in wet seasons, for the areas close to the target region (i.e., VMD), negative U-wind anomaly may increase the contribution evaporation (La Niña phase) and positive anomaly may decrease the contribution evaporation (El Niño phase). In the areas far away from the VMD, negative U-wind anomaly will reduce the amount of contribution evaporation. The possible explanation for this interesting phenomenon is that moisture in distant areas requires higher wind speeds to be transported to the target area, while close to the VMD, lower wind speeds may increase the possibility of precipitation formation because the duration of moisture residence in the VMD increased.

In addition to the influence of humidity and wind speed in the external area on the amount of contribution evaporation, the land-atmosphere interactions within the local area (VMD) can also affect the recycling ratio of moisture from the external area (i.e., Ratio_external in Figure 4.11), which is the ratio of contribution evaporation from the external area (CE_external) to the total amount of moisture from the external area. For example, lower soil moisture may decrease the likelihood of precipitation accompanied by drier atmospheric conditions (reduced evapotranspiration) (Seneviratne et al., 2010; Asharaf et al., 2012; Miralles et al., 2019). Ford et al. (2015) pointed out that morning soil moisture correlates very well with the changes in convective available potential energy (CAPE), which affects the initiation of convective precipitation. Hence, the relationship between local atmospheric conditions (vapor pressure deficit, VPD, CAPE and TCW_local) and the Ratio_external was explored (Figure 4.11). The VPD and TCW_local were utilized to represent surface and vertical integral atmosphere conditions, respectively, and CAPE was selected to illustrate the impact of convection on precipitation. In the dry season, the Ratio_external is highest while the atmosphere is unstable (CAPE > 70th percentile, the average is 750.1 J/kg) and wet (TCW_local > 70th percentile, the average is 49.1 kg/m^2). The effect of surface atmospheric conditions (i.e., VPD) on the Ratio_external seems not strong in the dry season. On the other hand, the Ratio_external is highest while the CAPE is lower than the 30th percentile (average is 782.4 J/kg) in the wet season. This result is consistent with Dong et al. (2019), who pointed out that larger values of CAPE do not imply higher precipitation. The explanation for this phenomenon is that the conversion of CAPE into kinetic energy becomes less efficient at larger values of CAPE, which affects condensation rates and precipitation. It is evident that, in both dry and wet seasons, atmospheric instability (i.e., CAPE) affects the Ratio_external, which typically reaches its maximum when CAPE values are approximately 750-780 J/kg.

Moisture transport and recycling are complex processes influenced by a variety of atmospheric and climatic factors. The findings presented in previous sections emphasized the significant roles of external atmospheric conditions, such as humidity and wind speed, in shaping these processes. In this section, the influence of the ENSO on moisture transport was further investigated, with a focus on its modulation of external atmospheric variables, including humidity and wind speed, which, in turn, affect precipitation dynamics in the VMD region. Additionally, local atmospheric conditions, such as CAPE, affect the atmospheric moisture recycling in the VMD.

Moreover, the potential impacts of other large-scale oscillations, including the Indian Ocean Dipole and the Madden-Julian Oscillation, on moisture transport and precipitation dynamics in the VMD warrant further investigation. Exploring the interactions among these large-scale atmospheric forcings and their effects on regional hydrological processes could enhance our understanding of the complex mechanisms governing moisture transport and recycling, thereby improving predictive models of precipitation variability in the VMD region.

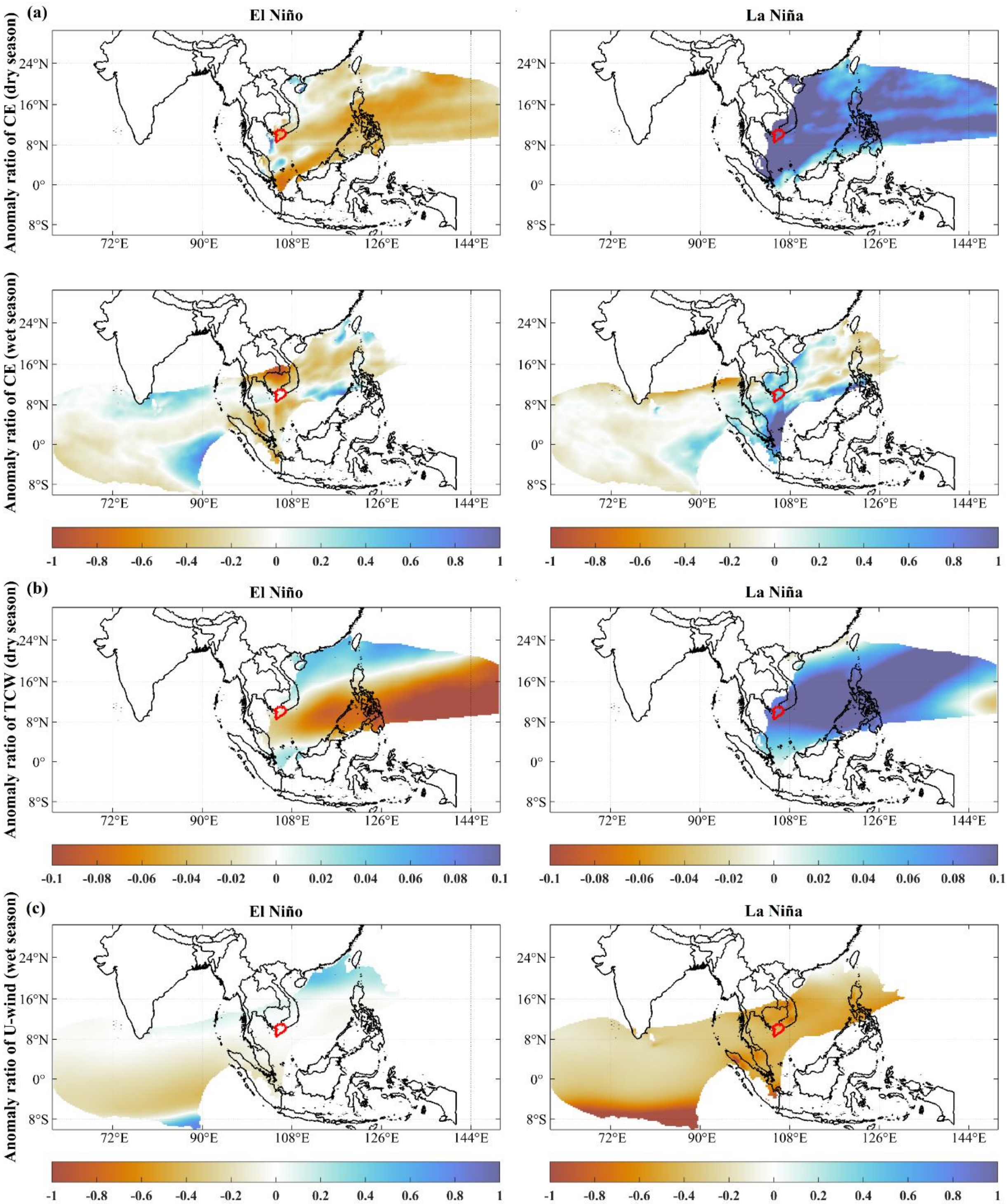

Figure 4.10. (a) The ratio between the anomalies in the contribution evaporation in the El Niño and La Niña phases and the contribution evaporation in the normal dry or wet seasons; The anomaly ratio between the (b) TCW and (c) U-wind in the phases of El Niño and La Niña and that in the normal dry or wet seasons.

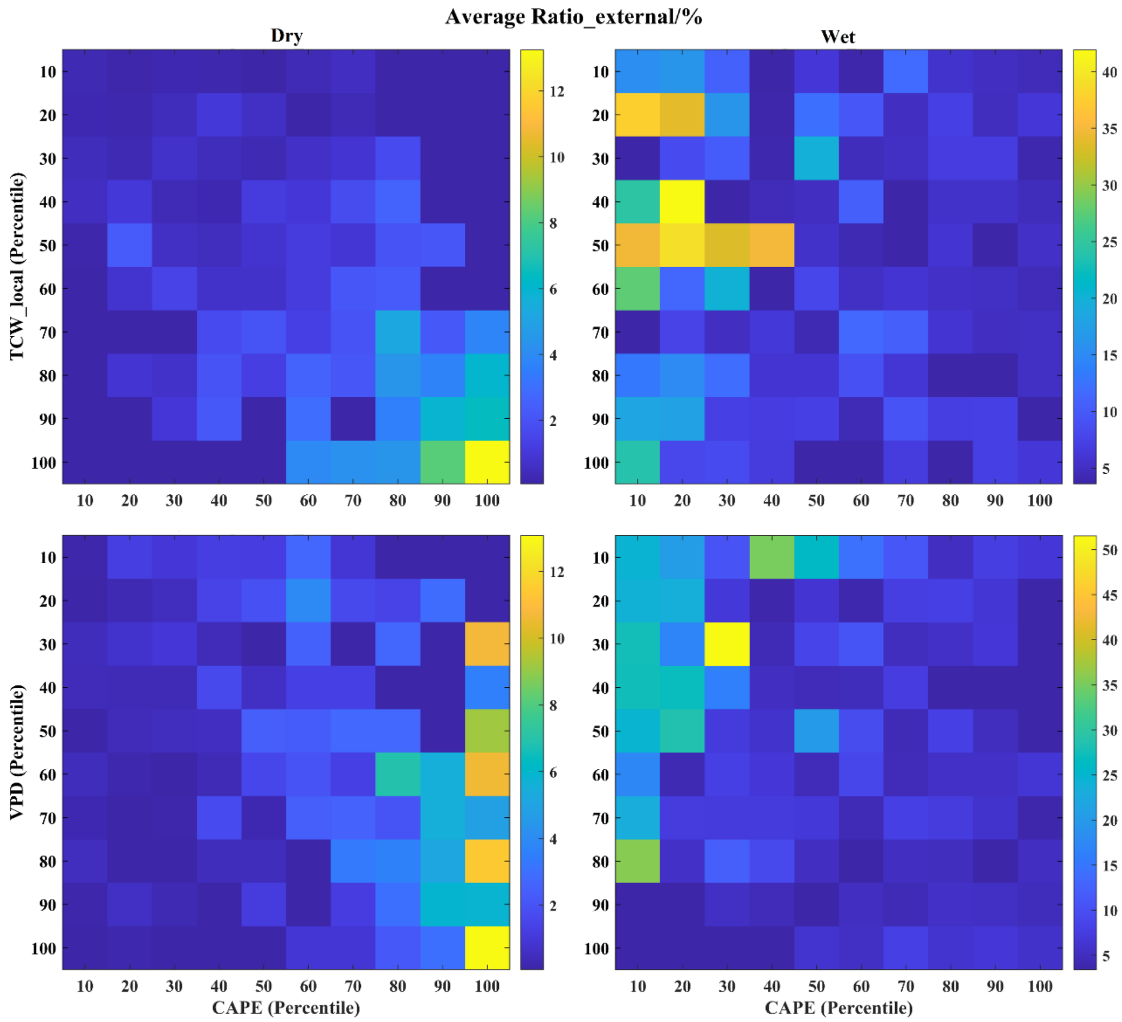

Figure 4.11. Averaged Ratio_external in each percentile bin between TCW_local/VPD and CAPE in the dry and wet seasons.

4.4 Discussion

4.4.1 Effects of external humidity and wind on the anomalies of contribution evaporation

In Section 4.3.2, the causality analyses revealed that humidity and wind speed in the upwind area are the two primary drivers of contribution evaporation and precipitation in the VMD during dry and wet seasons, respectively. Subsequently, in Section 4.3.3, the anomalies of these two variables were calculated to evaluate their roles in the reduction of moisture transport and precipitation in the VMD. However, this research didn't quantify the extent to which the contribution evaporation was affected by either specific humidity or U-wind in the upwind area. Yang et al. (2023) found that dryness was caused by the differences in horizontal wind convergence by comparing the moisture flux differences between control and experimental simulations. Similarly, Benedict et al. (2021) and Guan et al. (2022) decomposed the integrated vertical moisture flux into two components: dynamic (wind speed dominant) and thermodynamic (specific humidity dominant). It can be explained as follows (Seager et al., 2010; Li et al., 2013; Zhang, C. et al., 2017):

$$\int_{p_0}^{p} qudp = \int_{p_0}^{p} q_c u_c dp + \int_{p_0}^{p} q_c u_a dp + \int_{p_0}^{p} q_a u_c dp + \int_{p_0}^{p} q_a u_a dp \qquad (4.6)$$

where q and u represent specific humidity and U-component of wind respectively, q_c and q_a are the climatological average of specific humidity and its anomaly relative to the climatological average, u_c and u_a are the climatological U-wind and its anomaly relative to the climatology. $\int_{p_0}^{p} q_c u_a dp$ is the dynamic component that anomaly of moisture transport caused by the anomaly of U-wind, $\int_{p_0}^{p} q_a u_c dp$ is the thermodynamic component that anomaly of moisture transport caused by the specific humidity anomaly. Benedict et al. (2021) and Guan et al. (2022) found that moisture transport anomalies over the target study areas during the drought periods were controlled by dynamic processes. However, wind speed is a vector rather than a scalar (e.g., a positive or negative value means eastward or westward wind for U-wind, respectively). Therefore, the decomposition of moisture flux into thermodynamic and dynamic components may not fully capture the influence of humidity and wind speed over the upwind area on the amount of contribution evaporation. To explain this clearly, the 2020 dry season is presented as an example in Figure 4.12. Although humidity exhibited a negative anomaly over the most part of precipitationshed in the 2020 dry season (Figure 4.8), however, the thermodynamic component showed both positive and negative anomalies due to the contrasting climatological U-wind directions in the lower (westward) and upper (eastward) parts of the precipitationshed (Figure 4.12). Therefore, the anomaly in the thermodynamic component still cannot quantitatively reflect the effect of humidity on moisture transport. In addition, in dry and wet seasons, positive or negative anomalies of wind speed can have different effects on moisture transportation, which in turn influences precipitation in the target area. For example, the negative anomaly of U-wind in the 2020 dry season (Figure 4.8) accelerated the moisture transport to the VMD, while the similar negative U-wind anomaly in the 2019 wet season hindered the moisture transport and then amplified drought conditions in the VMD. In this study, therefore, in addition to the analyses of causal effects of humidity and wind speed on the VMD precipitation, only the anomalies of TCW and U-wind were considered to analyze their effect on the recent two severe droughts. However, it is still difficult to quantify the anomaly of contribution evaporation caused by external humidity and wind speed, respectively. It is mainly because the moisture tracking model was based on the atmospheric moisture budget equation (Van der Ent et al., 2014; Zhang, 2020; Guan et al., 2022). The amount of precipitation is hard to estimate under the conditions of climatological wind or humidity. Thus, it is difficult to directly track contribution evaporation with climatological wind or humidity. Separating the effects of wind and specific humidity on the amount of tracked contribution evaporation quantitatively would be an interesting topic in the future.

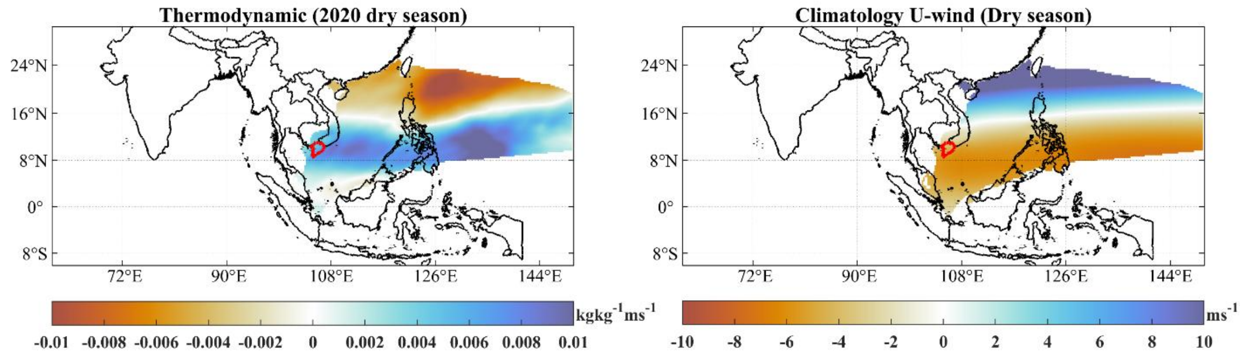

Figure 4.12. Distribution of thermodynamic components in the 2020 dry season and climatology of U-wind in the dry season.

4.5 Conclusions

To better understand precipitation moisture sources for the VMD during droughts, the moisture tracking model named WAM-2layers was applied in this study to identify and characterize the VMD precipitationsheds. To determine the dominant factors during the moisture transport process, the algorithms of PCMCI+ and CCM were also introduced to generate the causality networks. In addition, two recent record-breaking drought events were comprehensively analyzed as case studies. Through the full investigation for the VMD, some major findings were summarized below:

1. The precipitationshed was influenced by the seasonal northeast monsoon in dry seasons and southwest monsoon in wet seasons, respectively. The moisture from the external area contributes 60.4%-93.3% of the total contribution evaporation while the moisture from the local area accounts for 1.2%-27.1%. The recycling ratio of local total evaporation is lower in dry seasons than wet seasons.
2. The causality network is different for dry and wet seasons. In dry seasons, TCW_external influences the amount of contribution evaporation in the external area

and affects the amount of precipitation in the VMD. Wind speed affects the CE_external in wet seasons. Due to the drier condition in the upwind area in dry seasons, TCW is the factor that constrained the amount of CE_external. On the contrary, with sufficient water vapor (TCW) in the upwind area in wet seasons, the anomaly of wind speed dominates the amount of water vapor flux. The local atmospheric conditions (i.e., TCW_local, VPD and CAPE) may also have effects on the recycling ratio of moisture from the external area. The ENSO phenomenon affects the external atmospheric conditions and may impact the moisture transport and recycling for the VMD.

3. For the drought event in 2015-2016, the contribution evaporation reduced by around 40% relative to the climatology (5.74 mm and 79.61 mm for the local and external areas, respectively) in the 2015 dry season. The drought condition was relieved slightly in the latter part of the wet season after a dry start of the earlier part of the wet season. The dry season in 2016 was the worst drought period during the severe drought, with a reduction of around 80% in moisture from both external and local areas. For the drought in 2019-2020, the tracked evaporation even shows a negative anomaly in the wet season. The reduced moisture transport in the 2016 dry season was mainly caused by the anomalies of both humidity and wind speed, while the negative anomaly of moisture sources in the 2020 dry season was dominated by humidity only. In the 2019 wet season, the wind speed anomaly led to the reduction in moisture transport for the VMD.

The application of causality inference algorithms in this study substantially clarifies the understanding of the physical mechanisms underlying atmospheric moisture transport. By identifying the dominant factors influencing moisture transport at various stages, these algorithms provide valuable insights into the complex interactions between atmospheric variables. This deeper understanding of the physical processes governing moisture transport not only elucidates the mechanisms driving regional hydrological variability but also serves as a robust foundation for improving predictive capabilities of droughts in the VMD. For example, we can prioritize and pay attention to those dominant factors in drought prediction. Furthermore, with the indication and consideration of ENSO phenomena and external atmospheric conditions (i.e., humidity and wind speed), we can make more accurate drought predictions over a longer lead time. Future work could integrate causality inference with advanced climate models and explore its application across different temporal and spatial scales to further refine predictions and enhance regional climate resilience.

Chapter 5. Deep Learning-Based Quantitative Analyses of Land-Atmosphere Interactions over the Vietnamese Mekong Delta

Highlights:

- The deep learning model can effectively capture the relative importance of key variables in the Land-Atmosphere interactions.
- The degree of atmospheric response to anomalies in land surface state like soil moisture and sensible heat was quantified.
- The decline in soil moisture and the rise in sensible heat would raise temperature and further increase drought probability in the future.

Keywords: Drought; Land-Atmosphere interactions; Deep learning; Climate change; Vietnamese Mekong Delta

This chapter is a reformatted version of a manuscript published in the *Science of the Total Environment*, which is available at https://doi.org/10.1016/j.scitotenv.2024.175119

Abstract

During the past several decades, the Vietnamese Mekong Delta (VMD) has experienced many severe droughts, resulting in significant impacts on both agriculture and aquaculture. In the evolution and intensification of droughts, local Land-Atmosphere (LA) interactions were considered to play a crucial role. It is critical to quantify the impact of LA variables on drought processes and severity within the water and energy balances (e.g., soil moisture-latent and sensible heat-precipitation). In this study, a deep learning model, named Long- and Short-term Time-series Network (LSTNet), was applied to simulate the LA interactions over the VMD. With the ERA5 data as modelling inputs, the role of each key variable (e.g., soil moisture, sensible and latent heat) in the LA interactions over the period of 2011-2020 was quantified, and the variations of their inter-relationships in the future period (2015-2099) were also investigated based on the CMIP6 data. The LSTNet model has demonstrated that the deep learning algorithm can effectively capture the relative importance of key variables in the LA interactions. It is crucial to evaluate the effects of soil moisture and sensible heat on the LA interactions, particularly in the dry periods when negative anomalies in soil moisture and sensible heat would significantly reduce the amount of precipitation. In addition, the decline in soil moisture and the rise in sensible heat are anticipated to further diminish precipitation in the future under the changing climate.

5.1 Introduction

Droughts, as recurring extreme climate events, have profound and far-reaching impacts, regardless of the types of droughts considered (e.g., meteorological, agricultural, hydrological, and socio-economic droughts) and indices that are proposed (Wilhite and Glantz, 1985; Mishra and Singh, 2010; Li et al., 2016; Miralles et al., 2019). Droughts not only endanger water and food security, but also threaten the sustainability of ecosystems (Doughty et al., 2015; Anderegg et al., 2015; Miralles et al., 2019). Under a changing climate, the severity and frequency of droughts are projected to increase globally or regionally (Dai, 2013; Pokhrel et al., 2021; Li Y. et al., 2021; Dong et al., 2022; Kang et al., 2022). Therefore, the deep understanding of the physical processes that drive droughts is of great significance for social and environmental sustainability. It is widely believed that droughts are associated with large-scale atmospheric circulation anomalies (e.g., ENSO) and terrestrial processes (i.e., Land-Atmosphere (LA) interactions), while the former is critical for the initiation of droughts and the latter is crucial in the evolution and intensification of droughts (Seneviratne et al., 2010; Seager and Hoerling, 2014; Miralles et al., 2019; Holgate et al., 2020). Chapter 4 highlighted that over 60% of the precipitation moisture in the VMD originates from external regions, underscoring the significant contribution of remote moisture sources to the regional hydrological cycle. However, the results in Section 4.3.4 also revealed that local atmospheric conditions, such as atmospheric instability (e.g., CAPE) and local atmospheric humidity, also play a pivotal role in modulating moisture recycling efficiency. These findings emphasize the critical importance of local LA interactions in influencing the development and intensification of droughts. Therefore, this Chapter mainly investigated the role of LA interactions during the dry periods over the VMD (Figure 5.1), which is the most productive region in terms of agriculture and aquaculture in Vietnam but has greatly suffered from severe droughts in the past decade (Nguyen 2017; United Nations Resident Coordinator in Viet Nam, 2016, 2020). Previous studies on LA interactions have primarily concentrated on processes within the water cycle, such as the relationships between soil moisture, evaporation, and precipitation. The decline of soil moisture would reduce evaporation, dry the air, and may further inhibit the formation of precipitation and increase the likelihood of droughts occurrence (Santanello et al., 2013; Zhou S. et al., 2019; Schumacher et al., 2022). On the other hand, the reduction in soil moisture can contribute to the elevated and amplified sensible heat and temperatures, potentially provoking the onset of heatwaves (Hirschi et al., 2011; Mueller and Seneviratne, 2012; Miralles et al., 2014; Geirinhas et al., 2021). Concerning the coupled effects of soil moisture and temperature (sensible heat), prior research has predominantly concentrated on the co-occurrence of droughts and heatwave events (Hao et al., 2018b; Schumacher et al., 2019). Few studies investigated the interactions within the coupled water and energy balances (soil moisture-sensible heat-precipitation) associated with anomalies in sensible heat and precipitation. Dirmeyer et al. (2021) pointed out dry soil would alter surface fluxes, dry atmosphere, and exacerbate the drought and heatwave over Northern Europe. However, it is currently unclear to what extent these variables affect the process and severity of drought.

Previously, to quantify the degree of atmospheric response to anomalies in land surface state, the Global Land-Atmosphere Coupling Experiment (GLACE) was initiated with 12 Atmospheric General Circulation Models (AGCMs), in other words, GLACE was designed to measure the LA coupling strength (Koster et al., 2004, 2006; Guo et al., 2006). Moreover, given that soil moisture plays a key role in the LA interactions and climate systems, within the framework of the Coupled Model Intercomparison Project Phase 5 (CMIP5), the GLACE-CMIP5 experiment was implemented to investigate the impacts of soil moisture on the long-term changes in climate for both historical and future scenarios (Seneviratne et al., 2013; Schwingshackl et al., 2018). In the framework of GLACE-CMIP5 experiment, a control simulation was conducted with the original fully coupled soil moisture-climate interactions, while the experimental simulation was configured with the prescribed soil moisture (i.e., expA simulation: prescribed as the seasonal cycle of 1971-2000 climatology; expB simulation: prescribed as the seasonal cycle over a 30-year running mean). The difference between the control and experimental simulations can help us to quantify the contributions of soil moisture variability on the long-term climate changes in the LA interactions. However, it specifically investigates how and to what extent the soil moisture variability affects the climate system (e.g., precipitation, temperature), but ignore the role of other important climate factors (e.g., sensible heat) in the system. To comprehensively explore the inter-relationships among these variables in the LA interactions and quantitatively analyze the relative role of each variable in the system, deep learning algorithms would be an alternative approach (Shen, 2018). Up to now, deep learning algorithms have been rapidly developed and widely used in geoscience fields (Reichstein et al., 2019; Yu and Ma, 2021), including the investigations of natural hazards and extreme climate events (Liu and Wu, 2016; Liu Y. et al., 2016; Sharma et al., 2017; Racah et al., 2017), spatial and temporal state prediction (e.g., precipitation nowcasting, Shi et al., 2015; Zaytar and Amrani, 2016; Zhang et al., 2023), Earth system modelling (Gao et al., 2022; Bi et al., 2023), and data assimilation, downscaling and blending (Vandal et al., 2018; Niu et al., 2020; Liu et al., 2022b). With the development and maturity of deep learning algorithms, utilizing them to simulate the Earth system has the advantage of lower computation time and higher accuracy compared to traditional numerical physical models (e.g., Weather Research and Forecasting, WRF model and Regional Climate Model, RegCM model). Moreover, deep learning algorithms are no longer the full black box as previously described in geoscience research. Instead, we can infer their physical meanings from thoughtfully designed neural networks (Toms et al., 2020). In this study, the Long- and Short-term Time-series Network (LSTNet) deep learning algorithm (Lai et al., 2018) was employed to simulate LA interactions and comprehensively investigate the roles of key variables, including soil moisture, sensible heat, and latent heat, during droughts and dry periods. LSTNet was selected due to its proven effectiveness in handling multivariate time series predictions, as demonstrated in numerous prior studies (Ouyang et al., 2019; Bai et al., 2022). Its capability to capture both long-term patterns and short-term dependencies makes it particularly well-suited for analyzing the complex dynamics of LA interactions.

This study aims to investigate the inter-relationships among variables in the LA interactions over the VMD. Firstly, compare the outputs from the LSTNet and RegCM models to explore if the deep learning algorithm could perform better than the regional climate model in the simulation of LA interactions and try to interpret the physical meanings of the neural network. Then, the reference and experimental outputs were obtained with the original and prescribed data, respectively, to isolate the effects of each key variable in the LA interactions. Finally, the future impacts of how and to what extent the key variables would interlink in the LA interactions were also investigated based on the LSTNet model and the CMIP6 data.

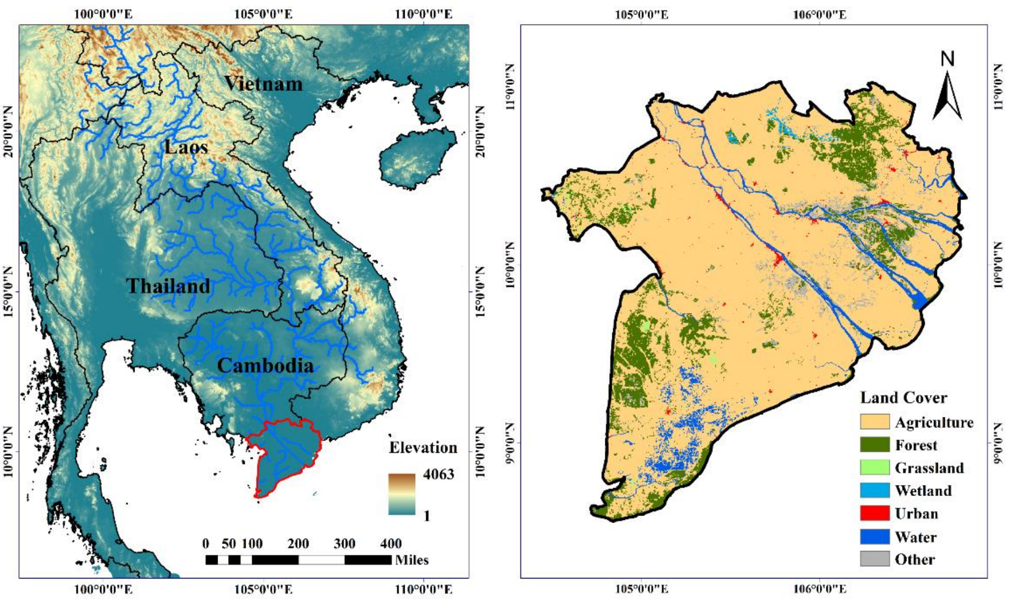

Figure 5.1. Geographical location and land cover types of Vietnamese Mekong Delta (VMD).

5.2 Materials and Methods

5.2.1 European Centre for Medium Range Weather Forecasts reanalysis version 5

ERA5 is the fifth-generation reanalysis product of the European Centre for Medium Range Weather Forecasts (ECMWF), covering the period from January 1940 to present (Hersbach et al., 2023a, 2023b). The ERA5 data used in this study include the variables at 16 pressure levels (i.e., specific humidity, temperature, geopotential, U- and V-component of wind) and those variables on the single level (i.e., total precipitation, surface sensible heat flux, surface latent heat flux, volumetric soil water layer 1 and 2, and sea surface temperature). These variables were selected due to the following two reasons: (1) variables at pressure levels could reflect the atmosphere conditions that may affect the LA interactions; and (2) the four land surface variables are the key components during the LA interactions. Here, the total precipitation, sensible heat and latent heat were downloaded at an interval of 1 hour, while other variables were derived at an interval of 6 hours. The spatial resolution of ERA5 is 0.25°× 0.25°. The data period used in this study was from 2010/12/01 to 2020/12/31. The sub-daily data in the first month (December 2010) were used as the model spin-up for the RegCM model (see Section 5.2.3). Table 5.1 summarizes the detailed information of different variables from the ERA5 and CMIP6 data in this study. As the inputs to the LSTNet model, the variables in the whole VMD were spatially averaged as multivariate time series. The ERA5 reanalysis data used in this study are freely available from the Copernicus Climate Change Service (C3S) Climate Data Store (https://cds.climate.copernicus.eu/cdsapp#!/home).

5.2.2 Coupled Model Intercomparison Project phase 6 dataset

To analyze the influential mechanisms in the LA interactions during drought events in the future scenarios, 6 CMIP6 models with 12 model simulations as shown in Table 5.2, were used in this study (Guo et al., 2018; John et al., 2018; Danabasoglu and Gokhan, 2019; Mizuta et al., 2019; Yukimoto et al., 2019; Lovato et al., 2021). The variables were derived from the CMIP6 models, including those variables at pressure levels (i.e., specific humidity, temperature, geopotential, U- and V-component of wind) and the variables on the single level (i.e., total precipitation, surface sensible heat flux, surface latent heat flux, moisture in upper portion of soil column). The variables were ensembled from the 12 model simulations. Several models have missing data in terms of variables at pressure levels. To ensure the continuity of the time series, data available over the VMD were spatially averaged. In instances where data for all grid cells over the VMD were missing, continuity was maintained by supplementing with data interpolated linearly from the adjacent pressure level. These models were selected because of the availability of those variables as indicated in Table 5.1. In addition, the chosen models offer a balanced compromise between temporal (1 day) and spatial resolution (25 and 100 km), which could reduce computational costs. The CMIP6 data are only used in the deep learning model to explore the future variations of these variables. The CMIP6 data derived from World Climate Research Programme (https://esgf-index1.ceda.ac.uk/projects/cmip6-ceda/).

Table 5.1. The variables derived from ERA5 and CMIP6 for setting up the RegCM and LSTNet models.

Data	Types	Variables	RegCM Inputs	LSTNet Inputs	Time span	Resolution
ERA5	Pressure levels	Specific humidity Temperature Geopotential U component of wind V component of wind	6 hours interval All the 37 pressure levels	6 hours interval 16 pressure levels (100, 200, 300, 400, 500, 600, 700, 800, 825, 850, 875, 900, 925, 950, 975 and 1000 hPa)	2010/12/01-2020/12/31	0.25°
	Single level	Total precipitation Surface sensible heat flux Surface latent heat flux	/	1 hour interval (sum to 6 hours interval)		
		Volumetric soil water layer 1 Volumetric soil water layer 2	/	6 hours interval (aggregate into 1 layer)		
		Sea surface temperature	6 hours interval	/		
CMIP6	Pressure levels	Specific humidity Temperature Geopotential U component of wind V component of wind	/	1 day interval 6 pressure levels (100, 250, 600, 600, 850, 1000 hpa)	2015/1/1-2099/12/31	See Table 5.2
	Single level	Total precipitation Surface sensible heat flux Surface latent heat flux Moisture in Upper Portion of Soil Column	/	1 day interval		

Table 5.2. 12 model ensembles from the CMIP6 data.

Experiment	Model	Ensemble	Resolution
ssp585 (RCP8.5)	NCAR CESM2	r4i1p1f1/r10i1p1f1/r11i1p1f1	100 km
	MRI ESM2	r1i1p1f1/r2i1p1f1/r3i1p1f1/r4i1p1f1/r5i1p1f1	
	CMCC ESM2	r1i1p1f1	
	GFDL CM4	r1i1p1f1	
	GFDL ESM4	r1i1p1f1	
highresSST	MRI AGCM3-2 H	r1i1p1f1	25 km

5.2.3 Regional Climate Model

The regional climate model can provide more detailed information attributable to its fine resolution, which considers the surface features as well as the meteorological processes (Nikulin et al., 2012). The model used for this study is RegCM Version 4.7.1 (https://github.com/ICTP/RegCM/releases) of the International Centre for Theoretical Physics (Giorgi et al., 2012; Gao et al., 2017), which incorporates the non-hydrostatic dynamical core of the fifth generation Penn State/NCAR Mesoscale Model (MM5) into RegCM4 (Giorgi et al., 2012). This integration involves specific modifications aimed at enhancing both the stability and adaptability of the model for long-term climate simulations (Coppola et al., 2021). The land surface scheme used in this study is CLM4.5 (Community Land Model version 4.5), and the land surface data for CLM4.5 is derived from http://clima-dods.ictp.it/regcm4/CLM45/. The lateral time-dependent boundary conditions use the exponential relaxation technique described in Giorgi et al. (1993). Convection is represented by the Emanuel scheme (Emanuel, 1991) and the atmospheric radiative transfer is computed using the RRTM radiation scheme, with the planetary boundary layer described by the nonlocal formulation of Holtslag et al. (1990). The RegCM runs at its standard configuration of 18 vertical sigma layers, with the model top level at 100 hPa. The ERA5 data is used to obtain the initial and lateral boundary conditions to drive the RegCM model.

5.2.4 Deep Learning-Based Model

The LSTNet (https://github.com/laiguokun/LSTNet) was developed for the prediction of multivariate time series, which consists of Convolution Neural Network (CNN), Recurrent Neural Network (RNN, i.e., Gated Recurrent Unit, GRU) and Fully Connected (FC) layers (Figure 5.2). The utilization of LSTNet is a good choice to simulate the interactions among the LA variables because it can extract short-term dependency patterns among the variables with CNN and discover long-term temporal patterns through the RNN layers (Lai et al., 2018), and the FC layer outputs the multivariate prediction. In addition, a novel recurrent-skip component was introduced in LSTNet, which leverages the periodic pattern of variables (Lai et al., 2018), for instance, the clear pattern on a daily basis in temperature. Specifically, recurrent-skip-links are added between the current hidden cell and those hidden cells with the same phase in the adjacent periods. Detailed information and structure of the neural network can be seen from Lai et al. (2018). In the training of all LSTNet models in this study, data segmentation was 70% for training, 20% for validation, and 10% for testing. Based on the ERA5 and CMIP6 data (Table 5.1), a total of 84 and 34 variables respectively were inputted into the model for training purposes. The simulation of all these variables constituted the training targets. Consequently, each variable contributes to the simulation of all these variables and their interactions. The parameters of the LSTNet model were determined based on the sensitivity test, which is illustrated as below:

According to the sensitivity test results in Figure 5.3, the model demonstrates enhanced performance with the CNN kernel size of 4, and the model is not sensitive to the length of the Window. Therefore, to balance the computational cost and model performance, the LSTNet model trained with ERA5 data (LSTNet_ERA5) utilized data spanning 14 days (6 hours per time step, 56 time steps in total) to predict the subsequent time step (lead time is 6 hours). As for the models trained with CIMP6 data (LSTNet_CMIP6), Consistent with LSTNet_ERA5, the models trained with CMIP6 data (LSTNet_CMIP6) also employed 14 days of data (1 day per time step, totaling 14 time steps) to predict the next 1-day variables. Although the LSTNet model exhibits improved performance with an increased number of hidden layers (Figure 5.3), however, considering the complexity of the model, 100 hidden layers were used in this study

In order to explore the internal structure of LSTNet and further analyze the importance of different variables in the simulation of LA interactions based on the neural network, the weights of CNN, GRU and FC layers were further investigated. To simplify the weight calculation for each time step in the GRU layer, the difference between GRU(X_0) and GRU(X_i) was used to represent the weights of each time step in the GRU layer, where $X_0 = [1, 1, 1, \ldots, 1]$, X_i is similar to X_0 but the i^{th} element of X_i is 0, e.g., $X_2 = [1, 0, 1, \ldots, 1]$, the length of X_0 and X_i represents the number of time steps of input data.

5.2.5 Simulation Experiments

For the historical analysis based on the ERA5 data, the LSTNet model (LSTNet_ERA5) was trained using the original multivariate time series. The LSTNet_ERA5 model was designed to simulate the interrelationships among various variables involved in LA interactions. Therefore, based on this model, different outputs can be obtained by prescribing the key variables (i.e., soil moisture, sensible heat, or latent heat) in the input data. Through the comparisons of outputs, the effects of the key variable on the other variables could be distinguished. For example, with the original multivariate time series as input, the reference output (REF_ERA5) was obtained based on the trained LSTNet_ERA5 model. Subsequently, by replacing soil moisture with its climatology time series in the input, an experimental output (EXP_ERA5_SM) was derived. Comparing the output of EXP_ERA5_SM with REF_ERA5 allowed us to distinguish the effects of soil moisture anomalies on other variables (e.g., precipitation, referenced in Figure 5.5(b)). Similarly, the experimental outputs for sensible heat (EXP_ERA5_SH) and latent heat (EXP_ERA5_LH) were obtained to isolate their respective effects during LA interactions. Figure 5.2(b) briefly illustrates the process of quantifying the effects of key variables on precipitation.

Parallel to the historical analysis, the LSTNet models (LSTNet_CMIP6) based on the CMIP6 data were also trained to examine future variations. In this study, the CMIP6 data was divided into four episodes: 2015-2034, 2035-2054, 2055-2074 and 2075-2099, and trained four separate LSTNet_CMIP6 models for each episode. Similar to the ERA5-based analysis, the effects of key variables on future LA interactions were evaluated by comparing the reference and experimental outputs derived from these models.

5.2.6 Evaluation Metrics

In this paper, the Pearson Correlation Coefficient (PCC), Root Mean Square Error (RMSE), and relative bias (BIAs) were used to measure the modelling performance. The equations for these metrics are shown in Table 5.3.

Table 5.3. Calculation of PCC, RMSE, BIAs, and ABIAs.

Metrics	Formula	Definitions
PCC	$\frac{\sum(O_i-\overline{O})\left(O_i'-\overline{O_i'}\right)}{\sqrt{\sum(O_i-\overline{O})^2}\sqrt{\sum\left(O_i'-\overline{O_i'}\right)^2}}$	O_i and O_i' represent two pairs of data for comparison
RMSE	$\sqrt{\frac{1}{n}\sum(O_i-O_i')^2}$	
BIAs	$\frac{\sum(P_i-R_i)}{\sum R_i}$	R_i and P_i represent input ERA5 data and models' predictions, respectively

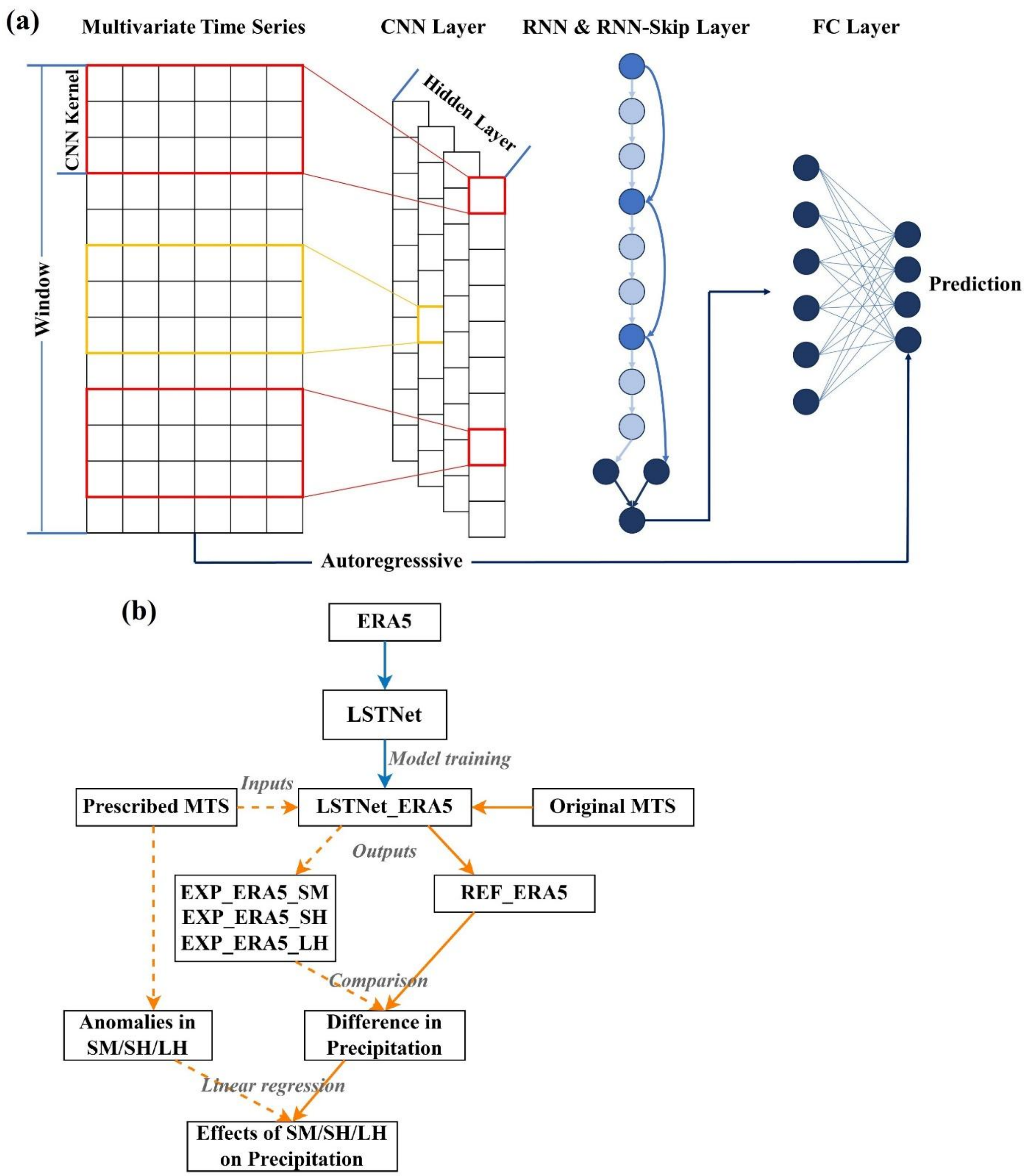

Figure 5.2. (a) Structure of the LSTNet model (the figure was reproduced based on Lai et al., 2018); (b) Flowchart of the simulation experiments for the analysis based on the ERA5 data.

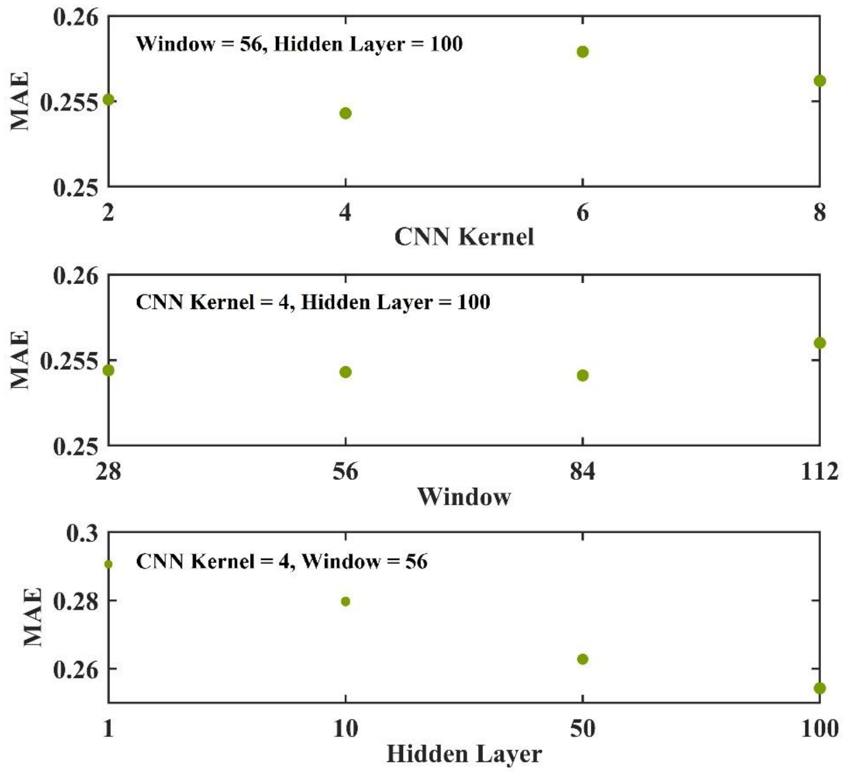

Figure 5.3. Sensitivity test of CNN Kernel, Window and Hidden Layer for the LSTNet_ERA5 model, the size of points represents the number of model parameters.

5.3 Results

5.3.1 Validation of LSTNet and RegCM modelling outputs with ERA5

The LSTNet and RegCM models can simulate different variables on the Earth's surface (e.g., soil moisture) and variables at multiple pressure levels (e.g., specific humidity). The four key variables (i.e., precipitation, soil moisture, sensible and latent heat) from the LSTNet and RegCM modelling outputs, which are related to the LA interactions, were validated with the ERA5 reanalysis data. To better evaluate the performance of LSTNet and RegCM models, here, the PCC and RMSE were calculated with anomalies in four variables from models' outputs and ERA5 data. As shown in Table 5.4 and Figure 5.4, based on the evaluation metrics, the LSTNet model performs better than the RegCM model in simulating four key variables. Relative to ERA5, the LSTNet model slightly underestimates the precipitation (BIAs = -0.082), especially in the dry seasons (i.e., December to April, BIAs = -0.258). In the dry seasons, the deviation of LSTNet predicted precipitation is greater and RMSE is lower than in wet seasons (i.e., June to October). This is mainly because of less precipitation in dry seasons. The RegCM model underestimates the precipitation even more than LSTNet, with the BIAs equal to -0.286, while the extreme precipitation of the RegCM model is even 1.6 times higher than the maximum ERA5 precipitation. The LSTNet performs very well in simulating soil moisture, which is basically consistent with ERA5 (PCC = 0.991, RMSE = 0.004 m^3/m^3, BIAs = -0.005). In general, the RegCM model underestimates soil moisture (BIAs = -0.123) compared with ERA5. As for the sensible and latent heat, the LSTNet also performs very well with the PCC higher than 0.78 and BIAs ranging from -0.04 to 0.006. The RegCM overestimates sensible heat in dry seasons (BIAs = 0.559) and underestimates it in wet seasons (BIAs = -0.980). The RegCM underestimates latent heat in both dry and wet seasons. Both LSTNet and RegCM perform better in simulating continuous and periodic variables (i.e., soil moisture, sensible and latent heat) than the variables with a discrete distribution (e.g., precipitation).

Table 5.4. Three evaluation metrics of LSTNet and RegCM modelling outputs with ERA5 data.

Metrics	Time	LSTNet				RegCM			
		Precipitation (mm)	Soil Moisture (m^3/m^3)	Sensible Heat (10^6 J/m^2)	Latent Heat (10^6 J/m^2)	Precipitation (mm)	Soil Moisture (m^3/m^3)	Sensible Heat (10^6 J/m^2)	Latent Heat (10^6 J/m^2)
PCC	Dry	0.703	0.993	0.878	0.803	0.526	0.778	0.727	0.415
	Wet	0.626	0.980	0.780	0.787	0.420	0.622	0.645	0.729
	All	0.632	0.991	0.858	0.792	0.443	0.668	0.704	0.586
RMSE	Dry	0.817	0.004	0.169	0.259	0.947	0.023	0.592	0.546
	Wet	1.427	0.004	0.135	0.328	2.002	0.020	0.273	0.380
	All	1.193	0.004	0.151	0.292	1.573	0.024	0.451	0.457
BIAs	Dry	-0.258	-0.005	-0.024	-0.010	-0.289	-0.134	0.559	-0.336
	Wet	-0.041	-0.006	-0.037	0.006	-0.248	-0.095	-0.980	-0.190
	All	-0.082	-0.005	-0.027	0.000	-0.286	-0.123	0.060	-0.243

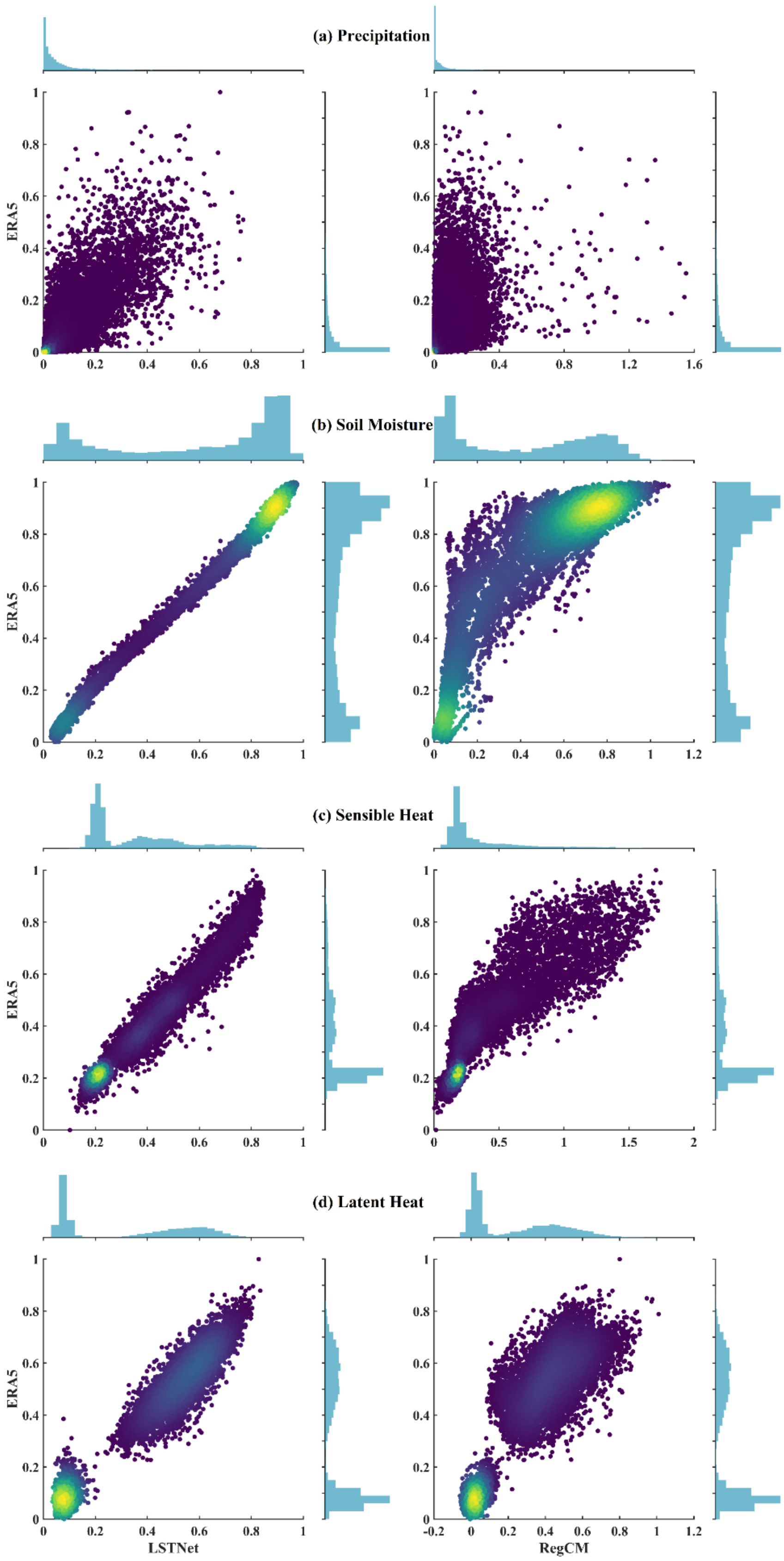

Figure 5.4. Scatterplots and histograms of normalized key variables: (a) precipitation, (b) soil moisture, (c) sensible heat, and (d) latent heat from the outputs of LSTNet and RegCM versus ERA5 over the VMD during the period of 2011-2020.

The above results reveal that the deep learning model outperforms the regional climate model in simulating the key variables of LA interactions. However, in order to quantify the effects of each variable in the LA interactions, it is crucial to explore the internal structure of the deep learning model to check whether it is a purely time series based black box or as an explainable framework that captures underlying physical characteristics. Therefore, as described in Section 5.2.4, the weights of the GRU layer were first analyzed. Generally, the GRU weights in the 8 time steps (*t*-8 to *t*-1 in Figure 5.5) that is close to the target time step *t* are much greater than those of earlier ones. From the perspective of the deep learning model, the prediction of variables is mainly based on the data from the last 8 time steps. As the time interval is 6 hours, it indicates that the LA interactions among these variables occurred within 2 days. Combined the weights of the CNN layer and FC layer, the weight of each variable in predicting four key variables (i.e., precipitation, soil moisture sensible and latent heat) was obtained. To better display the weights of these variables, for the variables on 16 pressure levels (e.g., specific humidity, temperature), their weights were aggregated into three levels as defined below: low level (900-1000 hpa), middle level (500-800 hpa), and high level (100-400 hpa). In addition, as shown in Figure 5.4, the normalized precipitation mainly concentrated on the range between 0 and 0.1, while the soil moisture is mainly between 0.8 and 1. To eliminate the impacts of different distributions of variables on the final weights, the long-term average of these variables was also included in the calculation of their relative importance. In Figure 5.6, the x-axis represents the abbreviations of variables: pr, sm, sh, lh, q, u, v, t, and z represent precipitation, soil moisture, sensible heat, latent heat, specific humidity, U and V component of wind, temperature, and geopotential, respectively, while subscripts h, m, and l represent high, middle and low-pressure levels, respectively. As illustrated in Figure 5.6, soil moisture, sensible and latent heat, wind speed and middle level specific humidity play an important role in the simulation of precipitation. Soil moisture, as a key regulator of surface energy fluxes (i.e., sensible and latent heat), exerts a substantial influence on atmospheric processes. Through latent heat flux, soil moisture affects atmospheric humidity, while sensible heat flux impacts air temperature and convection, both of which ultimately influence precipitation. The greater importance of sensible heat compared to latent heat in simulating precipitation may be attributed to the dominant role of convection in precipitation formation within this tropical regions. This finding aligns with the results presented in Section 4.3.4, which demonstrated that atmospheric instability (i.e., CAPE) significantly affects moisture recycling in the VMD.

Given that soil moisture is a continuous and periodic variable, antecedent soil moisture emerges as the most critical factor in simulating its future states in the LSTNet model. This is followed by the influence of latent heat, sensible heat, and precipitation. The simulation of latent heat is primarily influenced by soil moisture, low-level V-wind, and precipitation, while sensible heat is largely affected by low-level U-wind and soil moisture. The underlying reasons for latent heat being influenced by V-wind and sensible heat being affected by U-wind require further investigation in future studies. A possible explanation is that east-west circulation, governed by U-wind, drives surface temperature gradients and heat exchange, whereas north-south circulation, driven by V-wind, is crucial for transporting surface humid air masses.

Although the deep learning model employed in this study does not explicitly simulate LA interactions through predefined formulas or physical laws, the relative importance of variables and their weights, as illustrated in Figures 5.5 and 5.6, provides insights into the underlying physical processes. For instance, the persistence of soil moisture is shown to play a dominant role in the simulation of its own dynamics. These findings highlight the model's potential to uncover complex interactions among key variables, even in the absence of explicitly encoded physical principles.

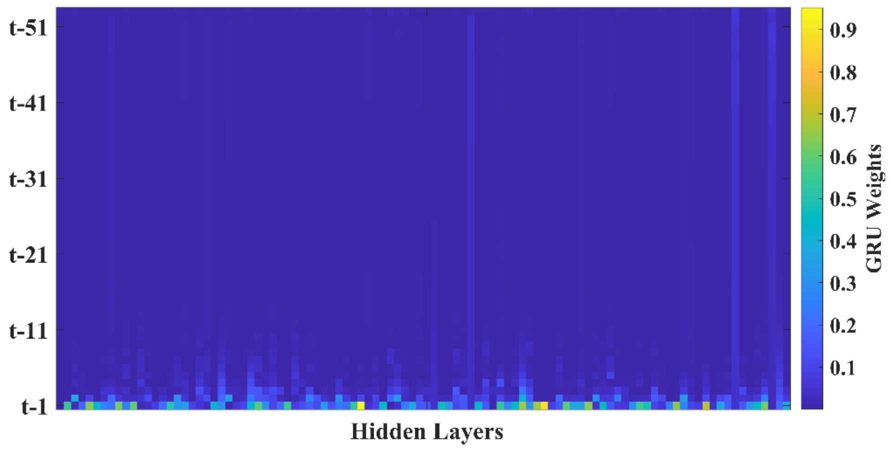

Figure 5.5. The GRU weights for different time steps and hidden layers in LSTNet.

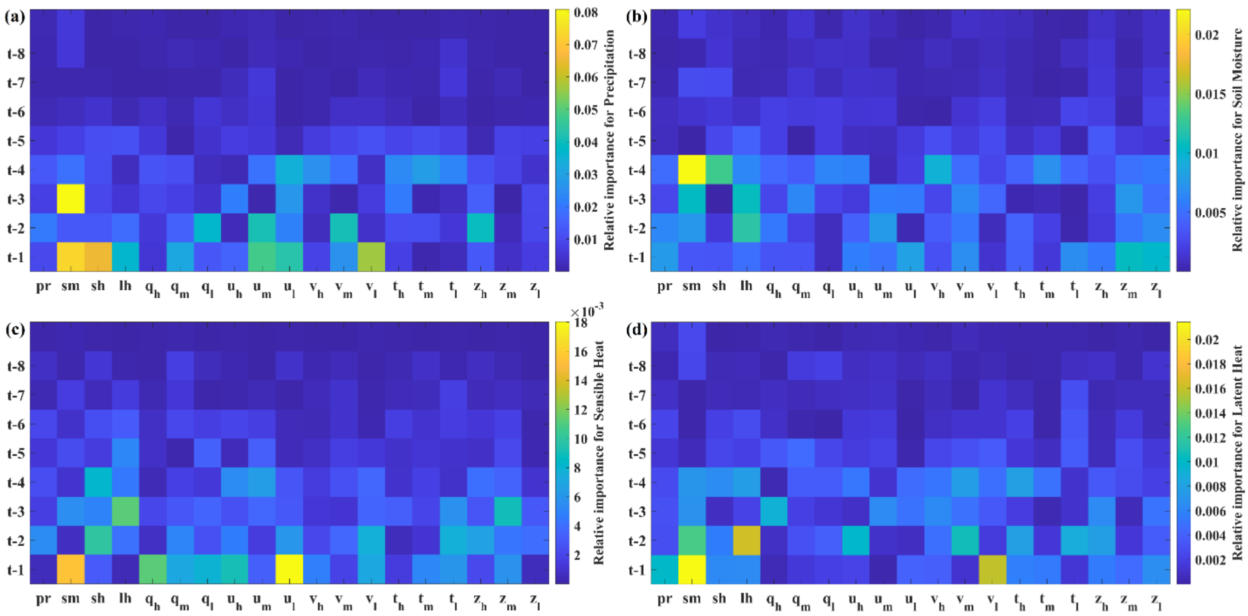

Figure 5.6. The relative importance of variables in the LSTNet simulation for the four key variables: (a) precipitation, (b) soil moisture, (c) sensible heat, and (d) latent heat.

5.3.2 The role of each key variable during the LA interactions

The results above have proved the efficiency of LSTNet in simulating LA interactions in the VMD, therefore, in the following sections, the effects of soil moisture, sensible and latent heat on precipitation were primarily analyzed and quantitated based on this deep learning model. First, the reference output (REF_ERA5) and experimental outputs (e.g., EXP_ERA5_SH) were obtained based on the original and prescribed input data respectively (Section 5.2.5). In the first row of Figure 5.7, the x-axis represents the percentage of anomaly in sensible heat between the original and prescribed input data (i.e., $(Original - Prescribed)/Original$), while the y-axis represents the difference in precipitation of two outputs, indicating the percentage of precipitation changes caused by sensible heat anomalies (i.e., $(REF\_ERA5 - EXP\_ERA5\_SH)/REF\_ERA5$). This study first analyzed the effects of anomaly in sensible heat, soil moisture and latent heat on precipitation, respectively. In this analysis, the dry and wet periods were defined with a soil moisture threshold of 0.25 m^3/m^3 (30% percentile). Figure 5.7(a) clearly shows that the impact of sensible heat anomaly on precipitation is significantly different during the dry and wet periods. The anomaly in sensible heat significantly affects precipitation in dry periods (PCC = -0.646, $p < 0.01$), and according to the regression slope in Figure 5.7(a), a 10% increase of sensible heat in the dry periods would reduce 1.53% precipitation. The anomaly of sensible heat can reduce precipitation by up to 20% in the dry periods. However, the effect of sensible heat on precipitation is not strong during the wet periods (Slope = -0.013, PCC = -0.358, $p < 0.01$). Soil moisture has a strong and significant influence on precipitation in both dry (Slope = 0.637, PCC = 0.861, $p < 0.01$) and wet periods (Slope = 0.517, PCC = 0.811, $p < 0.01$). As indicated in Figure 5.7(b), there is no significant difference between the dry and wet periods. The anomaly of soil moisture would reduce precipitation by up to 30% in the dry periods and increase precipitation by up to 20% in the wet periods. On the contrary, the effect of latent heat (Figure 5.7(c)) on precipitation is not strong in both dry and wet periods (Slope = 0.051, PCC = 0.168, $p < 0.01$). The increase and decrease in precipitation caused by the anomaly of latent heat are within 10% in the wet and dry periods.

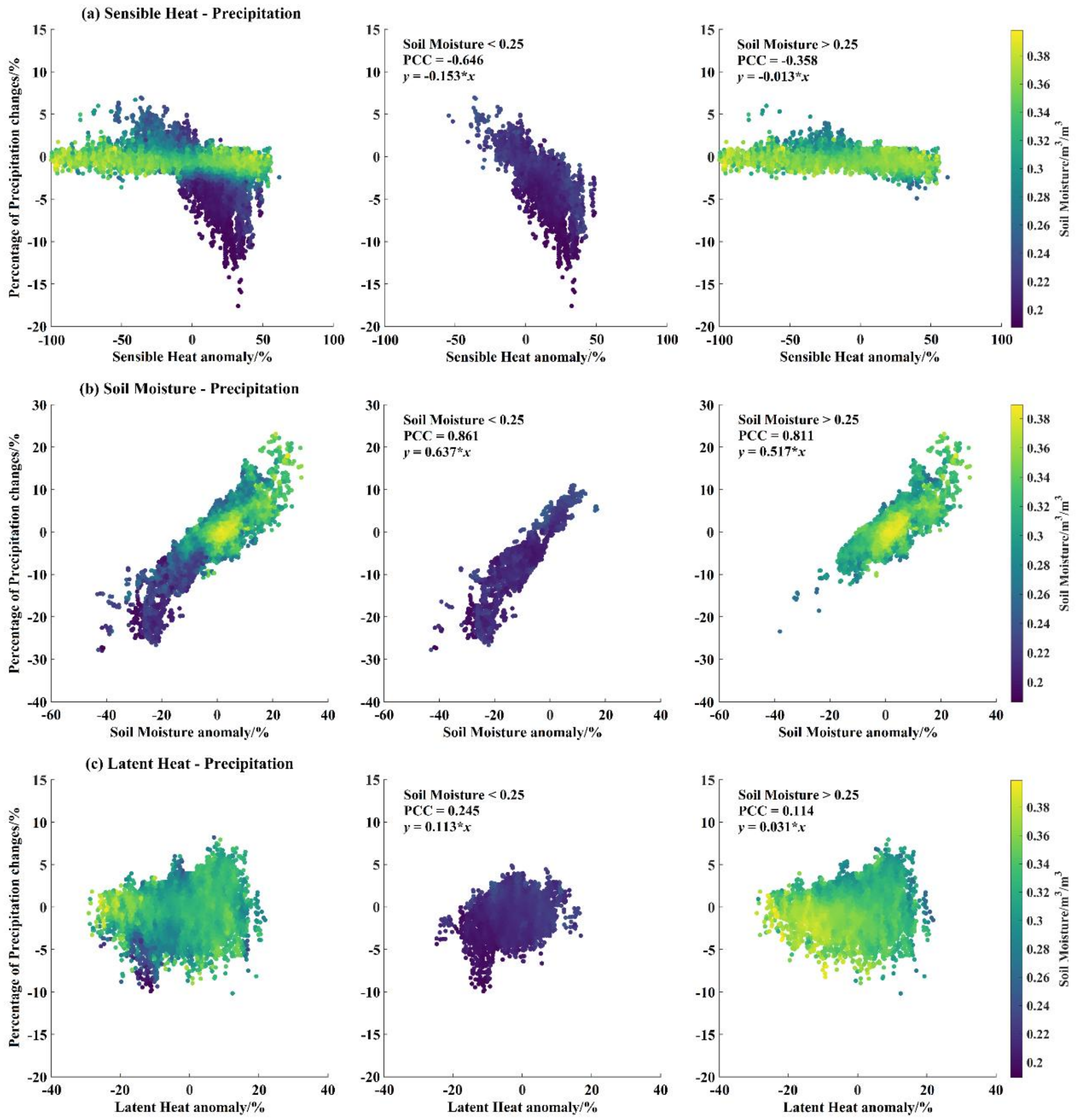

Figure 5.7. Scatterplots of the anomalies in: (a) sensible heat, (b) soil moisture and (c) latent heat versus the percentage of precipitation changes based on the reference and experimental outputs of LSTNet.

In addition, this study explored the effects of soil moisture on both sensible heat and latent heat, which are the important physical processes of LA interactions. As shown in Figure 5.8(a), the anomaly of soil moisture significantly affects sensible heat in both dry and wet periods, but the significance and strength of this effect vary between them. The 10% increase of soil moisture in the wet (dry) period would reduce 1.6% (0.71%) sensible heat. In the wet periods, the effect of soil moisture on sensible heat is approximately twice as strong as it does in the dry periods. The dry soil moisture may only increase sensible heat in the range of 0%-5%, while the anomaly of soil moisture may reduce 0%-15% of sensible heat in the wet periods. The effect of soil moisture on latent heat is significant and similar in the dry and wet periods as indicated in Figure 5.8(b). The changes in latent heat caused by the anomaly of soil moisture are within 3%. The quantitative analysis among these variables in the LA interactions is summarized in Figure 5.9.

The above results shows that precipitation may reduce while sensible heat increases in the dry periods. Furthermore, the severe drought event of 2015-2016 in the VMD was selected to explore the possible explanation and mechanism of this interesting phenomenon. Figure 5.10(a) illustrates the variations of normalized anomalies in sensible heat, convective inhibition (CIN), and precipitation changes. The background layer in Figure 5.10(a) represents the temperature anomaly at different pressure levels. The time span is from Jan. 2014 to Dec. 2016, which covers the drought event of 2015-2016. During the dry seasons of 2015 and 2016, Figure 5.10(a) clearly shows that the reduction (negative) in precipitation has a consistent agreement with the increased (positive) CIN and sensible heat. The positive anomalies in CIN and sensible heat correspond to the positive anomaly in temperature, especially in the low-level temperature (e.g. between 1000 and 925 hpa). In Figure 5.10(b), the high probability in the bottom-right bin indicates that the frequency of compound high sensible heat and high temperature is much higher than that expected if sensible heat and temperature were uncoupled. In Figure 5.10(c), each percentile bin shows the averaged CIN anomaly. The figure shows that the coupling of sensible heat and temperature is related to CIN. Therefore, the effects of sensible heat on precipitation observed in Figure 5.7(a) may relate to its coupling with temperature and convective processes.

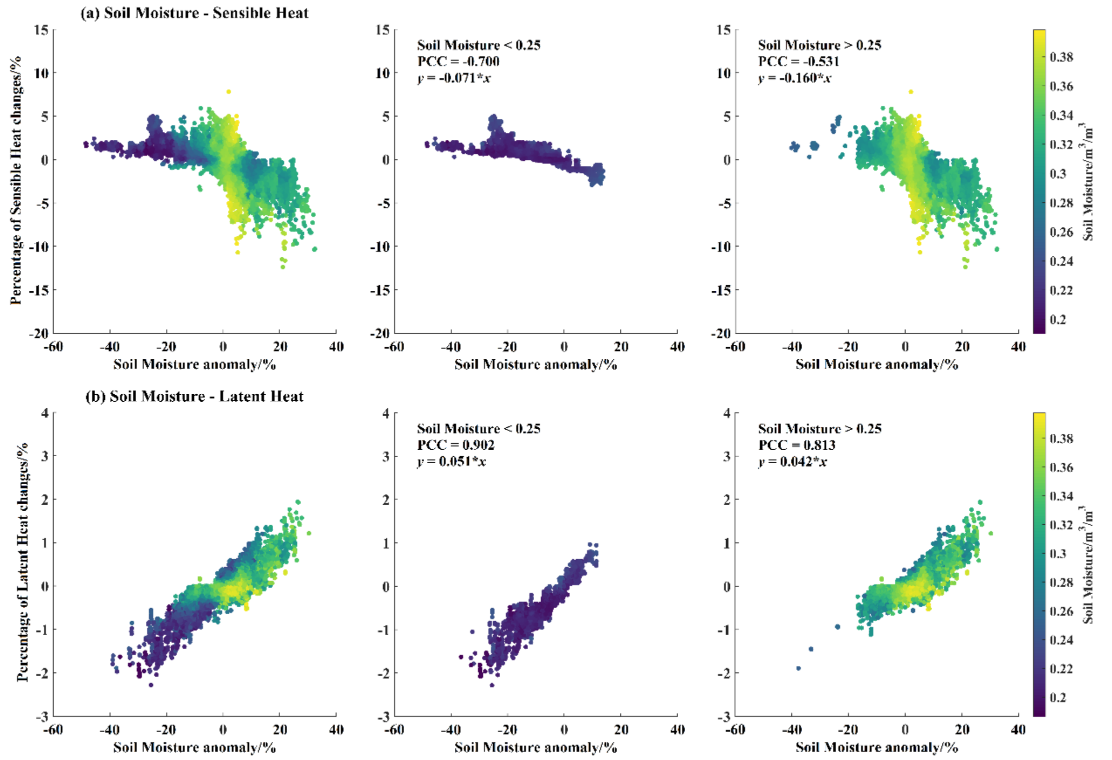

Figure 5.8. Scatterplots of the anomalies in soil moisture versus the percentage of (a) sensible heat and (b) latent heat changes based on the reference and experimental outputs of LSTNet.

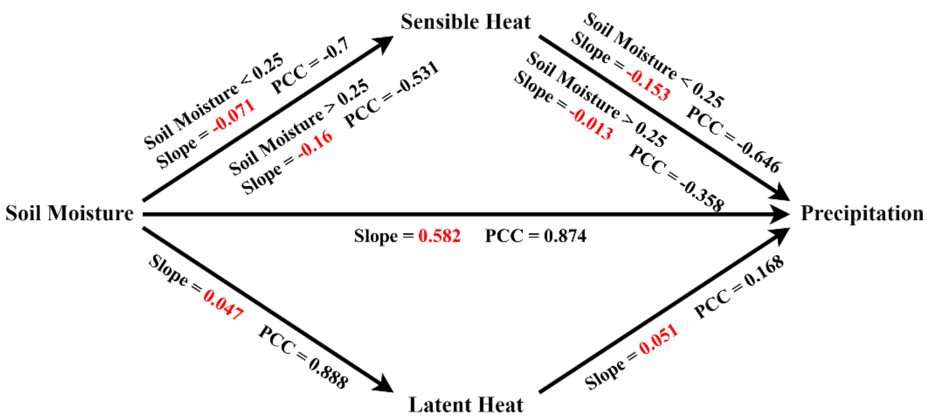

Figure 5.9. The summary of the quantitative analysis of the four key variables in the LA interactions.

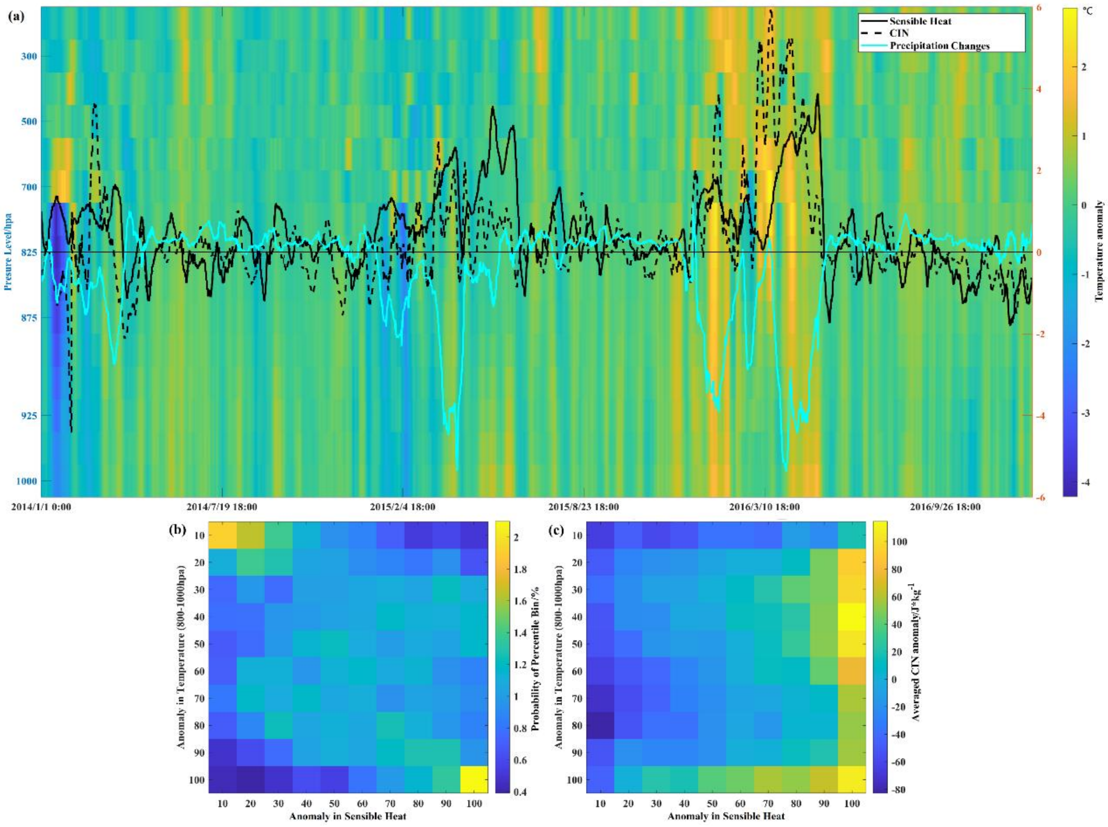

Figure 5.10. The effect of coupled temperature and sensible heat on the precipitation during drought: (a) the anomalies in sensible heat, CIN, precipitation changes and temperature at different pressure levels from 2014 to 2016; (b) mean probability of each percentile bin of averaged temperature anomaly (800-1000 hpa) and sensible heat anomaly; (c) averaged CIN anomaly of each percentile bin of temperature and sensible heat anomalies.

5.3.3 The role of each key variable during the future LA interactions

To explore future changes in the LA interactions, the variables from CMIP6 are used to develop the neural network (i.e., LSTNet). Here, the CMIP6 data were divided into four episodes: 2015-2034, 2035-2054, 2055-2074 and 2075-2099. Similar to the historical investigations in previous sections, this study mainly analyzed the effects of anomalies in sensible heat, soil moisture and latent heat on the precipitation simulation in the four episodes, respectively. Figure 5.11 shows that the impact of sensible heat anomaly on precipitation is quite different during the dry and wet periods. Here, the dry and wet periods were separated with soil moisture lower or higher than 21.5 kg/m^2 (30% percentile). The positive anomaly in sensible heat may reduce precipitation in the dry periods for all the four episodes. The 10% increase of sensible heat may reduce 1.14-3.76% of precipitation in the dry periods (Figure 5.12), which is consistent with the result from ERA5 during the period of 2011-2020. Episode

2 looks like a transition episode, with the lowest slope of -0.114 (PCC = -0.625, $p < 0.01$). From Figure 5.12, the sensible heat increases and the difference in precipitation decreases from episode 1 to episode 4. In episode 1, the averaged anomaly of sensible heat is negative which increases precipitation, and the averaged anomaly of sensible heat is positive which reduces precipitation in episodes 3 and 4. In the future, the rising sensible heat will reduce precipitation to some extent. Similar to the findings in Figure 5.10, the influence of future high sensible heat on the reduction of precipitation may relate to the increasing temperature under the context of global warming. On the other hand, the effect of sensible heat on precipitation is not strong during the wet periods for all the four episodes.

The anomaly of soil moisture significantly impacts precipitation, but in terms of its impact strength, there is a slight difference between the dry and wet periods (Figure 5.13). In the dry periods, the anomaly of soil moisture may reduce precipitation by up to 30% four episodes and these results are consistent with Section 5.3.2. From Figure 5.12, the slope between the soil moisture anomaly and precipitation changes in the first episode is 0.710, which is consistent with the slope of 0.637 in Section 5.3.2. Under the further development of global warming, the impact of soil moisture on precipitation will further aggravate in the dry periods (slopes ranging from 0.807 to 1.261 in the episodes 2 to 4). The soil moisture has a decrease trend from episode 1 to episode 4, which would cause the reduction of precipitation in the dry periods. In the wet periods, the differences in precipitation caused by the soil moisture anomaly mainly range from -20% to 20%, which is slightly lower than that in dry periods. The changes in precipitation caused by the latent heat anomaly mainly range from -5% to 5% in the dry periods of the first three episodes. In episode 4, the averaged anomaly of latent heat decreases a lot and leads to the reduction of precipitation (Figure 5.14). The variation of latent heat anomaly in the four episodes is not as strong as sensible heat and soil moisture (Figure 5.12).

Figures 5.15 and 5.16 display the effects of soil moisture anomaly on the simulation of sensible heat and latent heat respectively. The differences in sensible heat caused by soil moisture anomaly range from -5% to 5% in the dry periods of four episodes. With soil moisture may decrease in the future, the sensible heat may increase simultaneously (Figure 5.17). The slopes between the soil moisture anomaly and sensible heat difference range from -0.17 to -0.24 in the dry periods, which is lower than the results based on the ERA5 data (Slope = -0.071), but the same pattern is that the effect of soil moisture anomaly on sensible heat is stronger in the wet periods than that in the dry periods. Similarly, the latent heat differences caused by soil moisture anomaly range from -5% to 5% in the dry periods. The slopes between the soil moisture anomaly and latent heat difference ranges from 0.23 to 0.35 in dry periods of all four episodes (Figure 5.17). The effect of soil moisture anomaly on latent heat in the wet periods based on the CMIP6 data is smaller compared with dry periods (Figure 5.16). When soil moisture decreases in the future, the latent heat would decrease simultaneously (Figure 5.17).

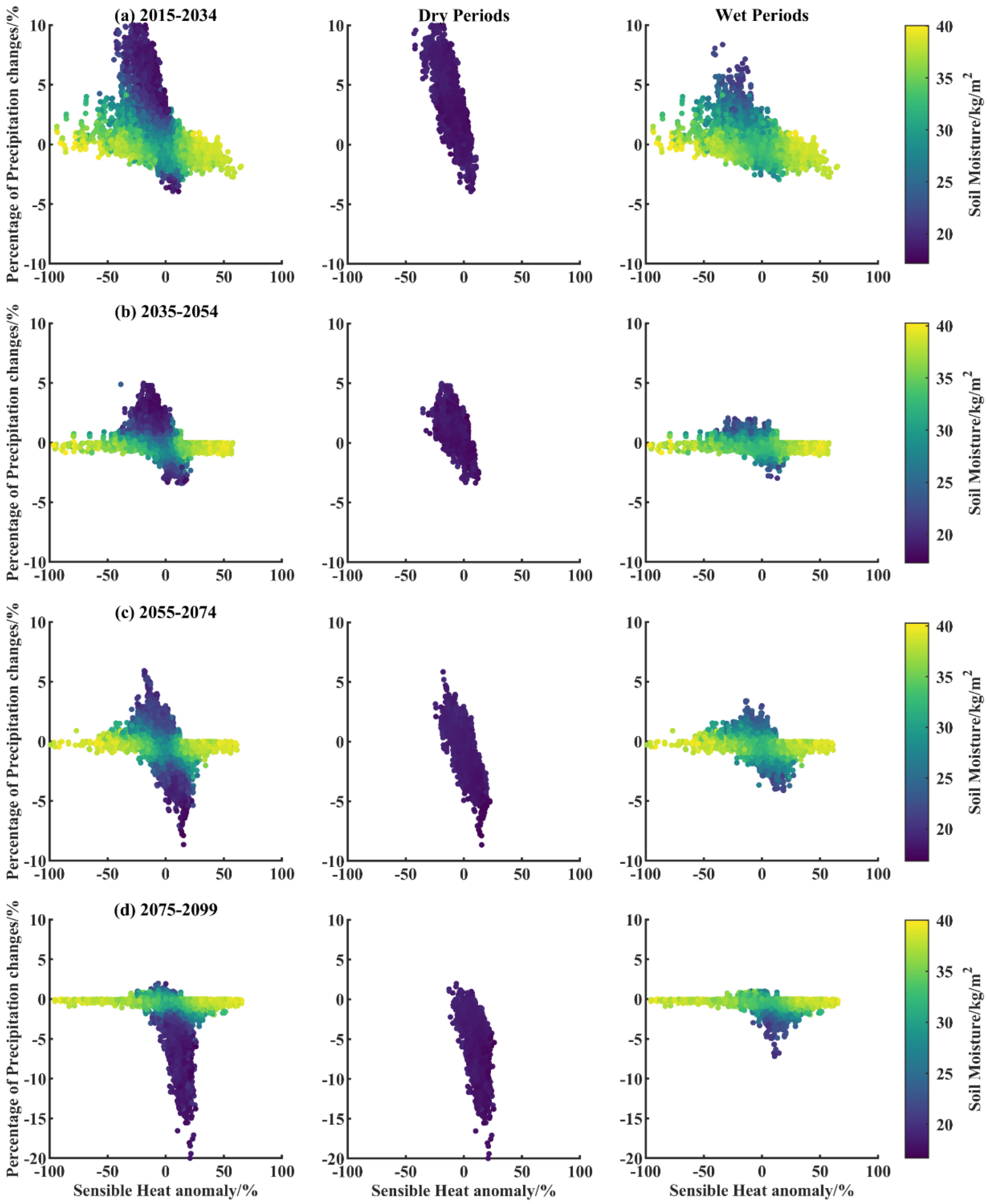

Figure 5.11. Scatterplots of the anomalies in sensible heat versus the percentage of precipitation changes during the four episodes: (a) 2015-2034, (b) 2035-2054, (c) 2055-2074 and (d) 2075-2099.

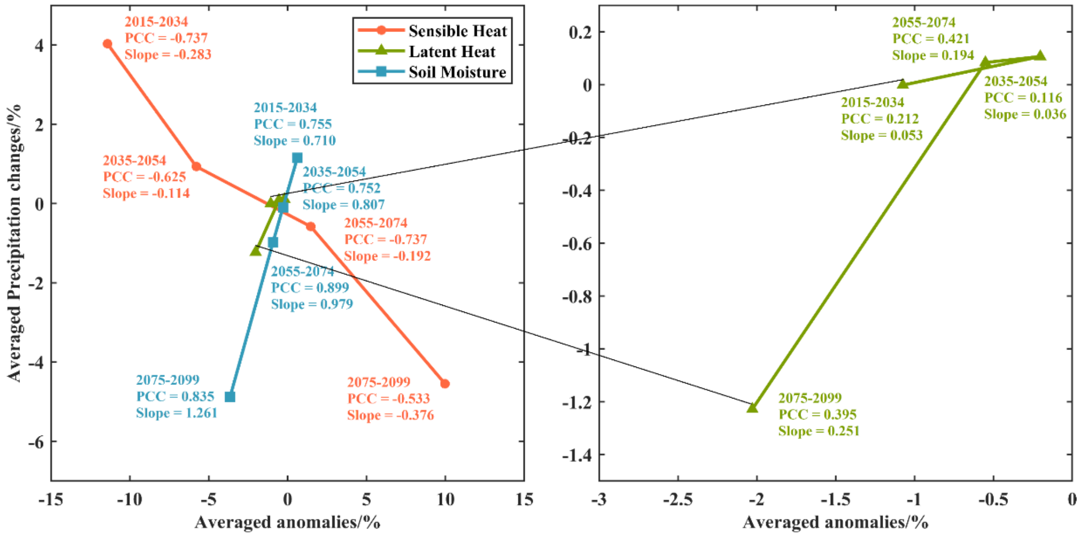

Figure 5.12. The variation of precipitation changes caused by the anomalies in sensible heat, latent heat, and soil moisture (average of dry periods points).

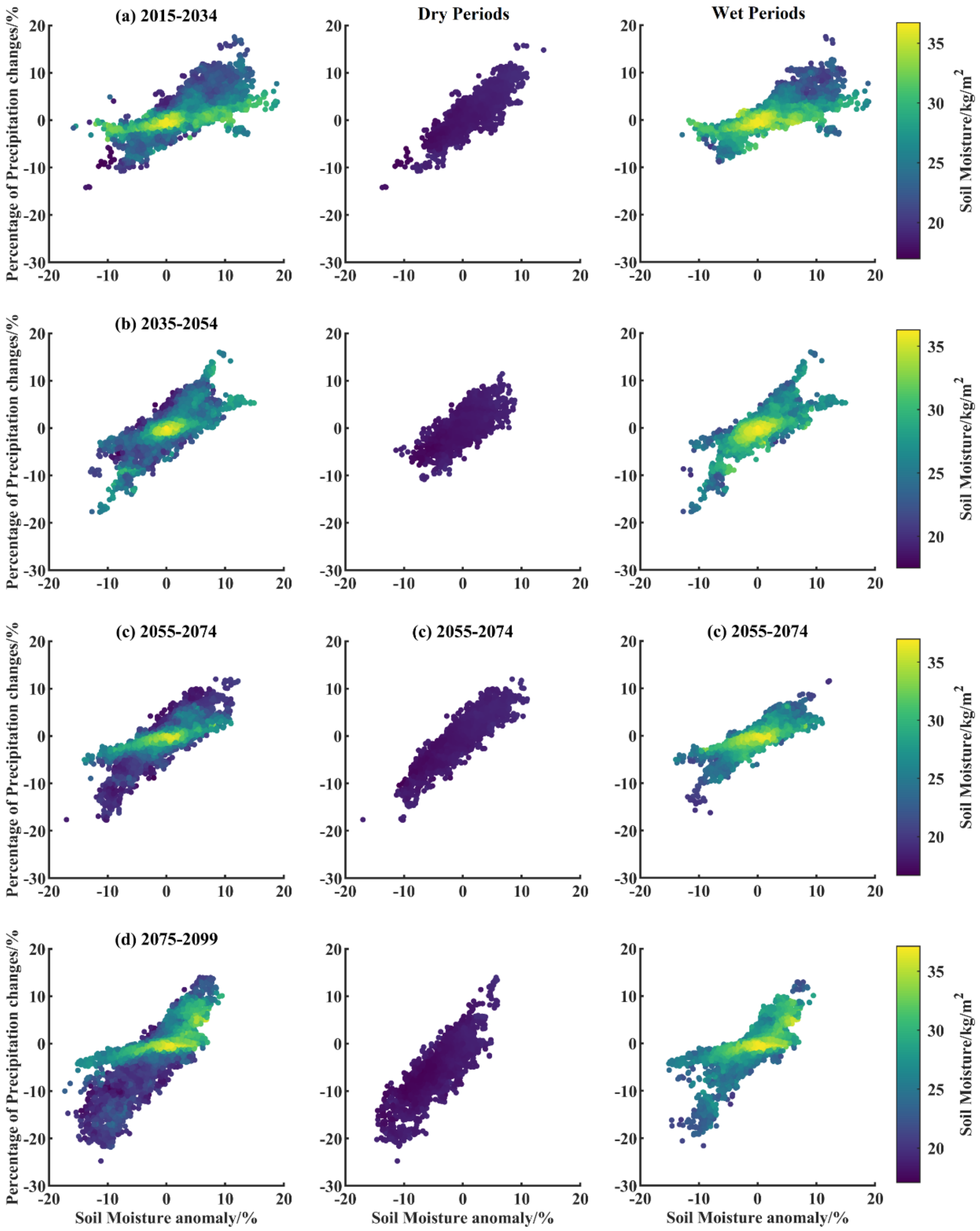

Figure 5.13. Scatterplots of the anomalies in soil moisture versus the percentage of precipitation changes during the four episodes: (a) 2015-2034, (b) 2035-2054, (c) 2055-2074 and (d) 2075-2099.

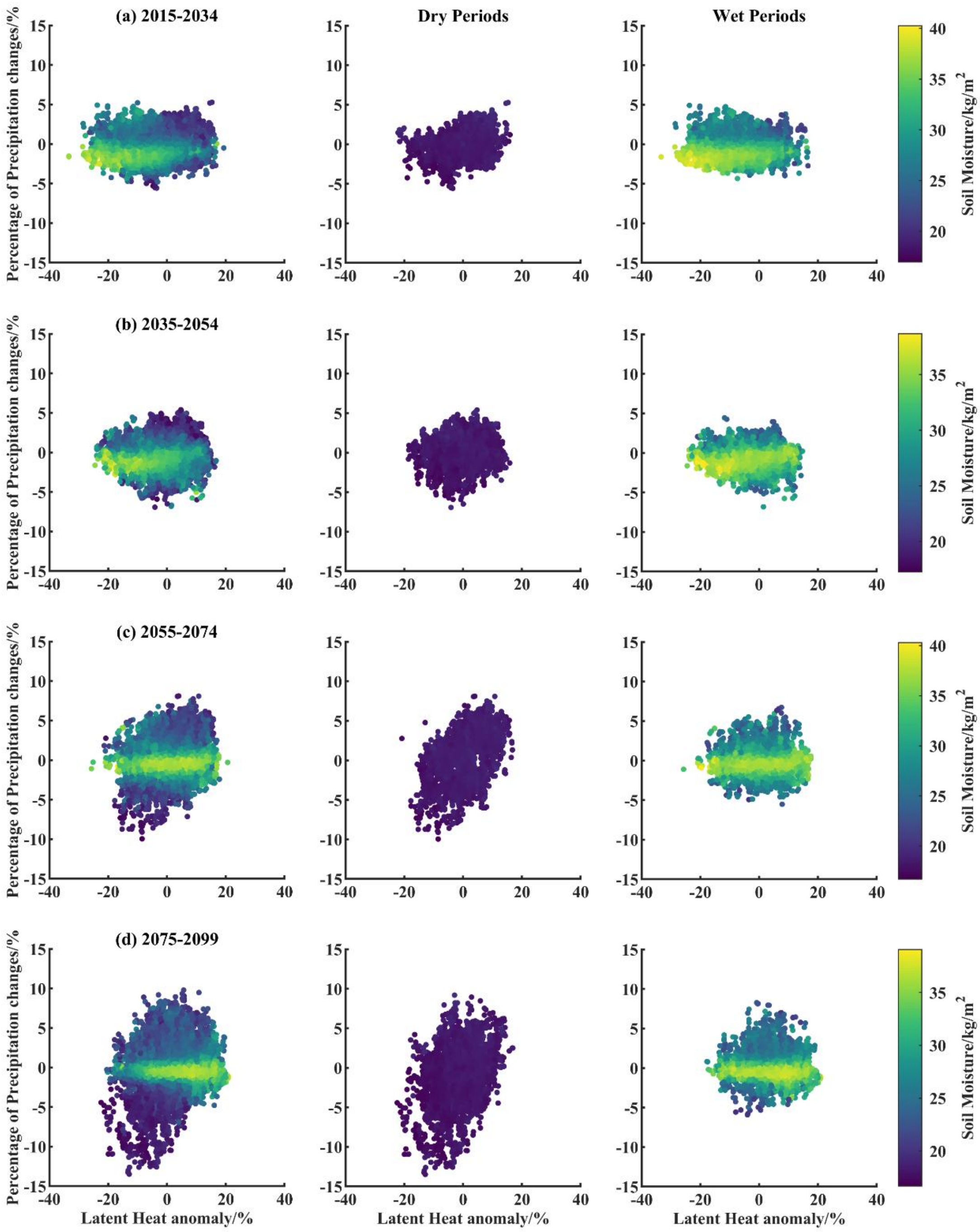

Figure 5.14. Scatterplots of the anomalies in latent heat versus the percentage of precipitation changes during the four episodes: (a) 2015-2034, (b) 2035-2054, (c) 2055-2074 and (d) 2075-2099.

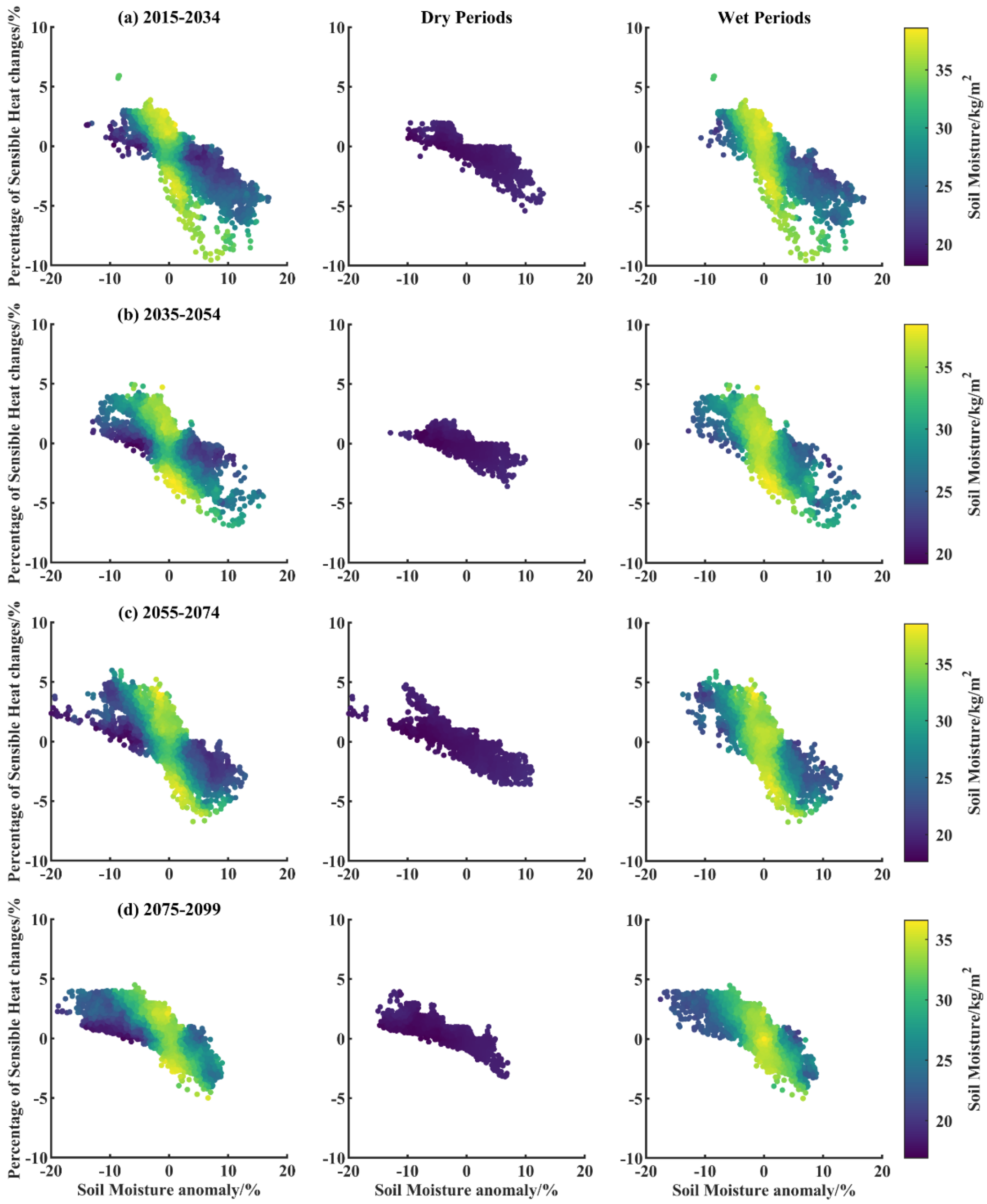

Figure 5.15. Scatterplots of the anomalies in soil moisture versus the percentage of sensible heat changes during the four episodes: (a) 2015-2034, (b) 2035-2054, (c) 2055-2074 and (d) 2075-2099.

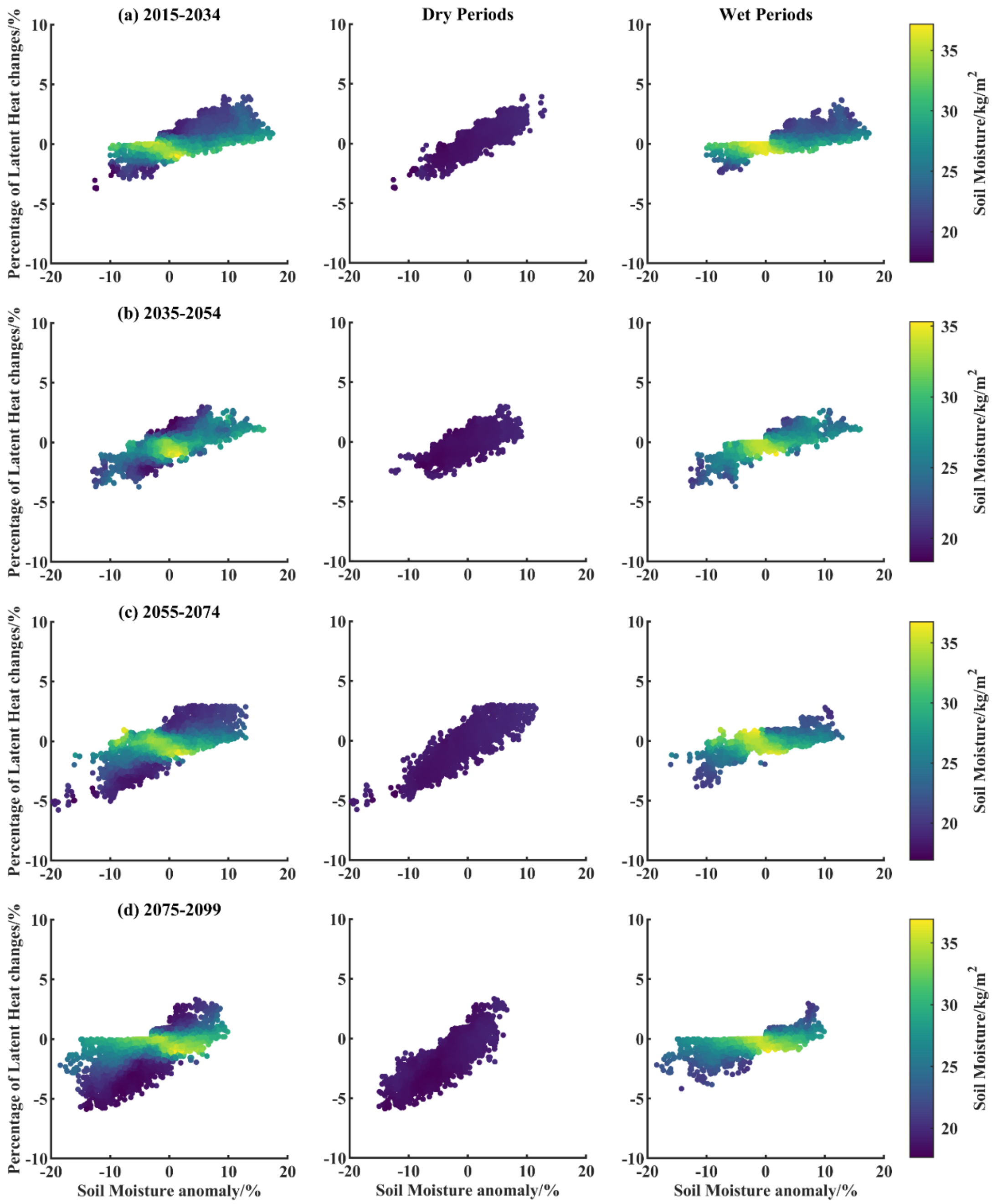

Figure 5.16. Scatterplots of the anomalies in soil moisture versus the percentage of latent heat changes during the four episodes: (a) 2015-2034, (b) 2035-2054, (c) 2055-2074 and (d) 2075-2099.

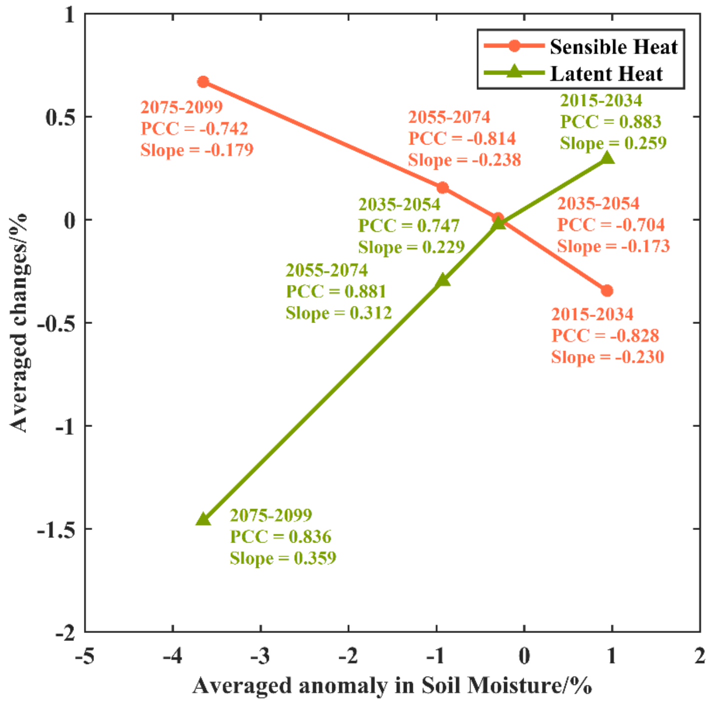

Figure 5.17. The variation of sensible and latent heat changes caused by the anomaly in soil moisture (average of dry periods points).

5.4 Discussion

5.4.1 Potential applicability of LSTNet neural network in simulating LA interactions in other climate regions

The LA interactions shall vary in different climate zones across the world, therefore, a global-scale investigation is valuable to explore the interrelationships among variables in the LA interactions with deep learning algorithms. By taking Central California (i.e., Central Valley) as an example, we analyzed the effects of key variables during the LA interactions and discussed the applicability of the LSTNet neural network in different regions. As one of the major agricultural regions in the United States, California experienced the most severe drought and heatwave in 2014, which caused 2.2 billion dollars economic loss and massive groundwater overdraft (AghaKouchak et al., 2014; Mann and Gleick, 2015; Seager et al., 2015). According to the climate projection, the drought situation will be more severe in the middle of this century and co-occurred with more extremely hot days in California (Ullrich et al., 2018). As mentioned in Section 5.3.2, the positive anomalies in sensible heat may enhance the deficiency of precipitation. Hence, an LSTNet neural network was trained for Central California to investigate the role of sensible heat in simulating precipitation in the future. Like findings in the VMD, the impact of sensible heat on precipitation is quite different during the dry and wet periods as shown in Figure 5.18 (with a threshold of 30% percentile in soil moisture, 17 kg/m^2). The slope between the sensible heat anomaly and precipitation difference in the dry periods of four episodes ranges from -0.041 to -0.056, which is slightly higher than that in the VMD. Also, the sensible heat will increase under the context of global warming and would reduce precipitation to some extent in California in the future (Figure 5.19). Figure 5.19 shows that the averaged anomalies of sensible heat in the dry periods present increased trends (from -3.1% to 3.7%) that caused the averaged difference in precipitation to decrease from 1.4% to -1.7%. The LSTNet neural network can effectively capture the role of key variables (e.g., sensible heat) in the LA interactions in different climate regions, which is important for understanding drought occurrence and development in different climate regions.

5.4.2 Cross validation with the GLACE-CMIP5 experiments

The GLACE-CMIP5 has been widely applied to analyze the effects of soil moisture on temperature and precipitation extremes (Lorenz et al., 2016) and on land carbon uptake (Green et al., 2019; Humphrey et al., 2021). Based on the GLACE-CMIP5 experiment, the soil moisture-atmosphere feedbacks would alleviate the decrease of water availability in drylands (Zhou S. et al., 2021) and exacerbate droughts and atmospheric aridity in the future (Berg et al., 2016; Zhou S. et al., 2019). This study modelled the LA interactions based on the LSTNet neural network and evaluated the effects of key variables in these interrelationships for the VMD and Central California. The LSTNet is a deep learning algorithm rather than a physical process-based model. Even though the LSTNet can capture the relationships among the variables very well, it would be of great significance to validate the LSTNet outputs with the GLACE-CMIP5 model. Firstly, we discussed how the soil moisture anomaly affects surface temperature at 1000 hpa in the VMD and Central California and compared them with the GLACE-CMIP5 results. The differences in temperature caused by the soil moisture anomaly range from -0.1 to 0.1 ℃ in the VMD and from -0.3 to 0.3 ℃ for Central California, respectively (Figure 5.20). The drier the soil moisture, the higher the temperature in both Central California and VMD, and the effect of soil moisture on temperature is much stronger in Central California (Slope = -1.368) than in the VMD (Slope = -0.350). The soil moisture trend may increase the temperature by approximately 0.5℃-1.5℃ in the VMD and Central California (difference between the expB and expA simulations, Lorenz et al., 2016), where the soil moisture shows a decreasing trend (Dirmeyer et al., 2013; Lu et al., 2019). Then, the dry trend of soil moisture may prolong consecutive dry days (daily precipitation < 1 mm, Lorenz et al., 2016). Both the LSTNet neural network and GLACE-CMIP5 experiments have indicated the same patterns: 1) the lower the soil moisture, the higher the temperature; 2) the lower the soil moisture, the less the precipitation (Figure 5.13). With the proposed development of LSTNet-based global simulations, further cross validation with GLACE-CMIP5 can be carried out beyond the VMD and Central California.

5.4.3 Uncertainty in the deep neural networks

Predictive uncertainty in neural networks arises from two primary sources: data uncertainty and model uncertainty. Data uncertainty encompasses variability in real-world situations, measurement system errors, and errors caused by unknown data. Model uncertainty includes errors in the neural network architecture specification and training procedures (Loquercio et al., 2020; Hüllermeier and Waegeman 2021; Gawlikowski et al., 2023). For example, in this study, uncertainty may stem from the inputted ERA5 and CMIP6 data, as well as from the configuration of the LSTNet neural network, including choices in batch size, optimizer, and learning rate. As illustrated by Hersbach et al. (2020), the globally averaged uncertainty (ensemble spread) in ERA5 decreases over time, with the lowest spread observed near the surface for variables such as zonal wind and specific humidity. Figure 5.21 presents that from 2011 to 2020, the uncertainty in precipitation, soil moisture, sensible and latent heat in the VMD remained generally low, generally below 1 mm, 0.02 m^3/m^3, 10^5 and 2×10^5 J/m^2, respectively. These variables exhibit clear seasonal patterns in uncertainty. The uncertainty of precipitation, sensible heat, and latent heat in wet seasons is greater than that in dry seasons, while the uncertainty of soil moisture is the opposite. The uncertainty of CMIP6 data mainly includes three different sources, namely internal variability, model uncertainty, and scenario uncertainty, with model uncertainty being the dominant source (Hawkins and Sutton, 2009; John et al., 2022; Wu, Y. et al., 2024). Additionally, the linear regression for interpolating missing data may also introduce uncertainty during the training of the neural network. On the other hand, quantifying model uncertainty in neural networks remains a significant challenge. Bayesian neural networks, which learn a distribution over weights, represent the state-of-the-art for estimating predictive uncertainty. However, they require substantial modifications to the standard training procedures and are computationally demanding compared to non-Bayesian networks (Gal and Ghahramani, 2016; Lakshminarayanan et al., 2017). Therefore, it is still valuable to investigate and quantify the errors and uncertainties derived from the inputted data and neural networks, which is of great significance for improving the performance and reliability of neural networks.

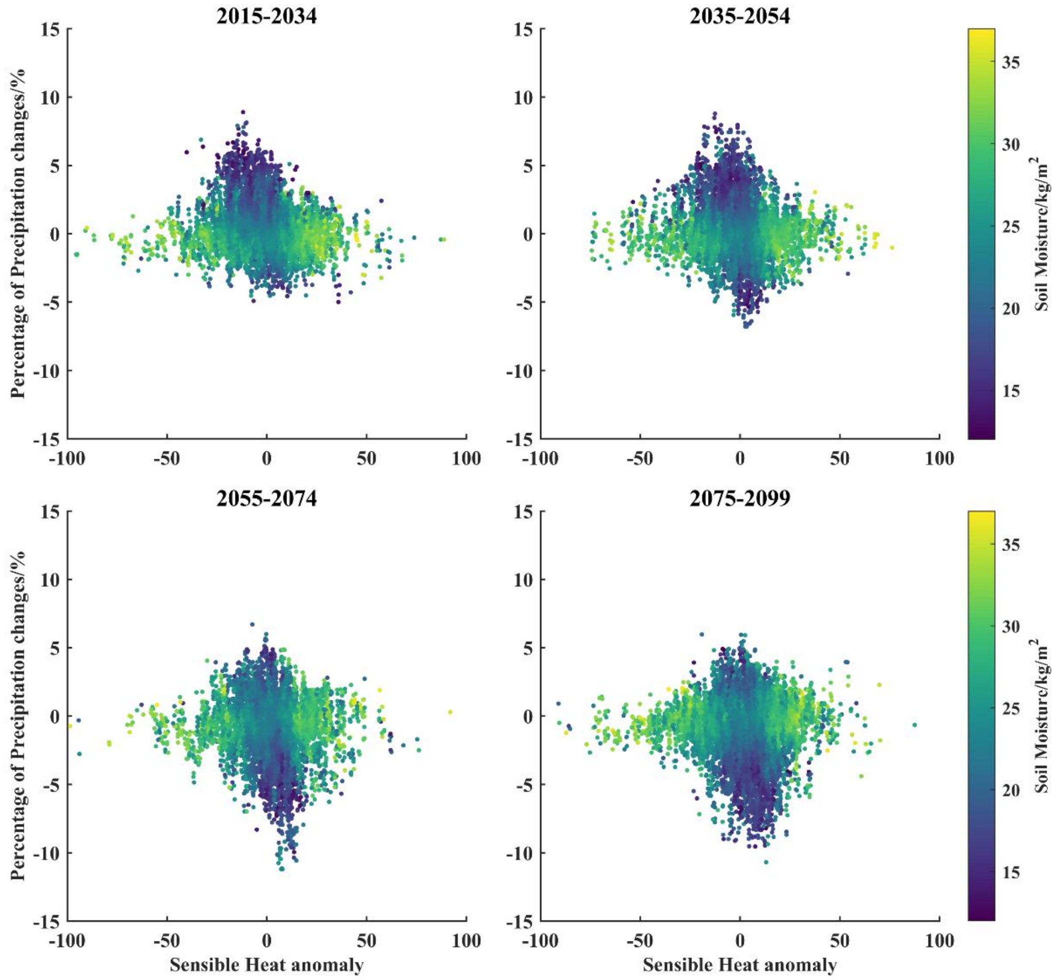

Figure 5.18. Scatterplots of the anomalies in sensible heat versus the percentage of precipitation changes during the four episodes in Central California.

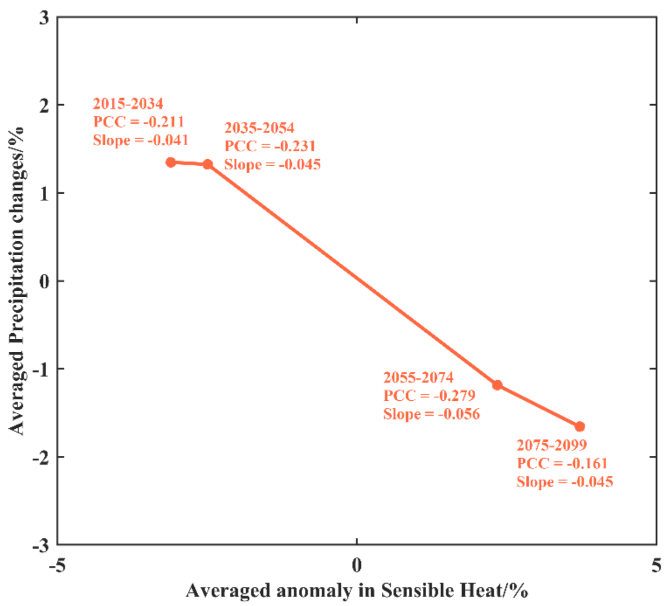

Figure 5.19. The variation of precipitation changes caused by the anomaly in sensible heat in central California (average of dry periods points).

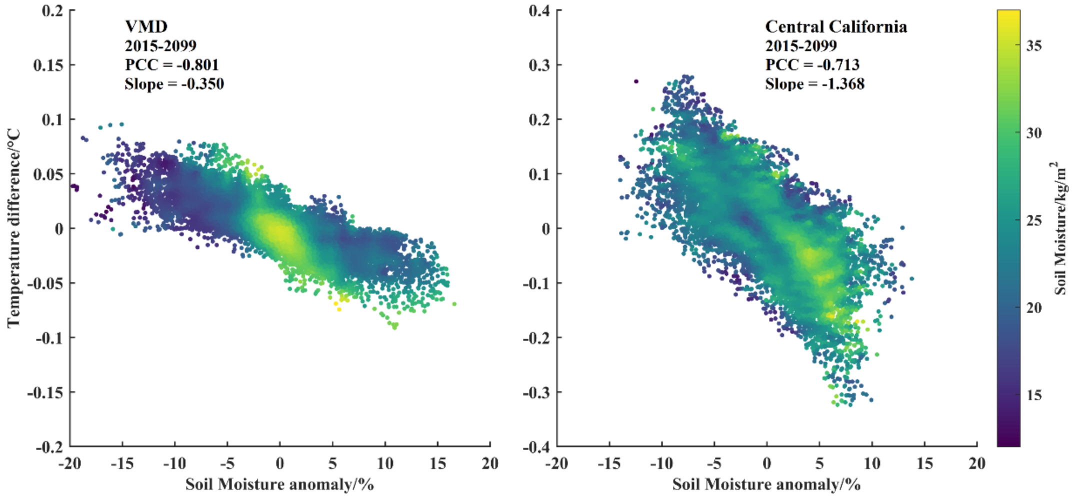

Figure 5.20. Scatterplots of the anomalies in soil moisture versus the temperature difference at 1000 hpa in the VMD and Central California.

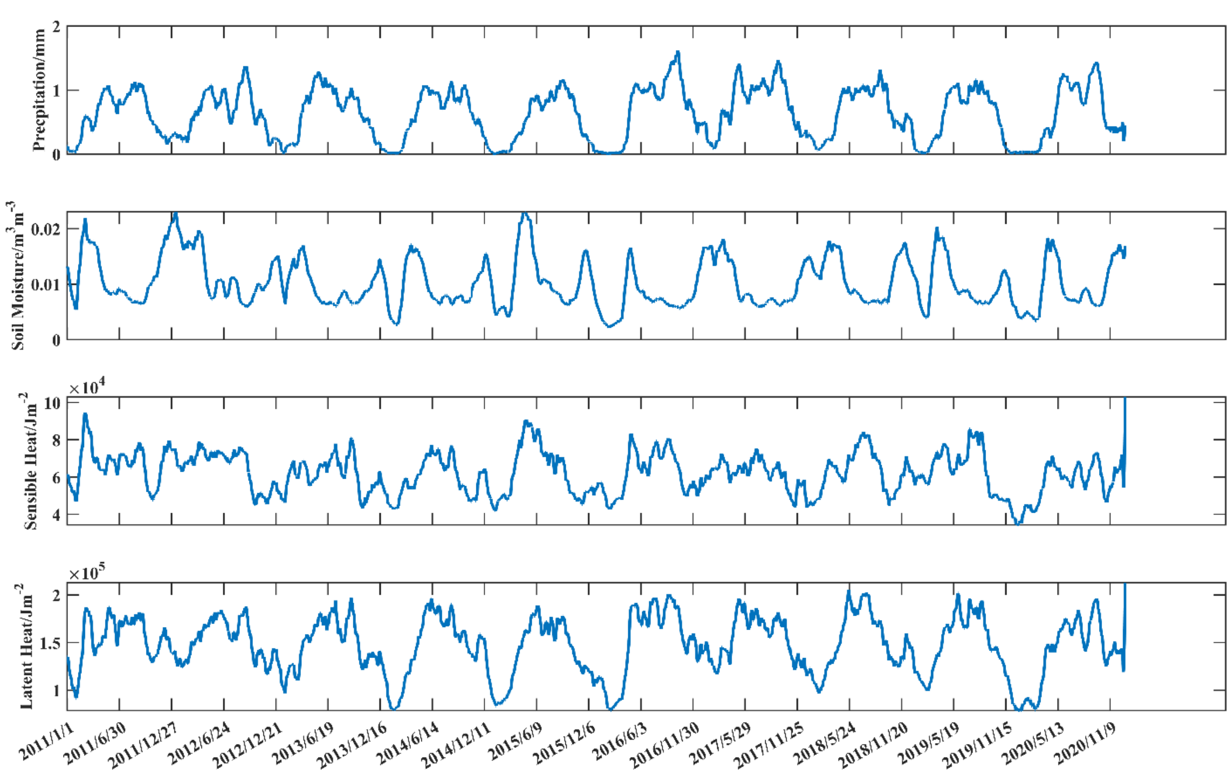

Figure 5.21. Time series of 30-day averaged ERA5 ensemble spread (uncertainty) from 2011 to 2020 for precipitation, soil moisture, sensible and latent heat in the VMD.

5.5 Conclusions

This study verified the applicability of deep learning algorithms (i.e., LSTNet) in the simulation of LA interactions. Then, the effects of each key variable (i.e., soil moisture, sensible and latent heat) were isolated in the LA interactions based on the reference and experimental outputs, which were implemented with the original and prescribed data, respectively. Using the CMIP6 data as the deep learning modelling inputs, how and to what extent the key variables would affect future LA interactions were further analyzed. Main findings are listed below:

1. Compared with the RegCM model, the LSTNet neural network performed better in simulating the four key variables during both dry and wet seasons. In the VMD, local interactions among those key variables occurred within 2 days. The weights and relative importance of LA variables indicate that the LSTNet algorithm uncovers complex interactions among key variables, even without explicitly physical principles.
2. The anomaly of sensible heat can reduce precipitation by up to 20% in dry periods, while in wet periods, the impact of sensible heat on precipitation is not as strong as that in the dry periods. Soil moisture has a great influence on precipitation in both dry and wet periods. The effects of sensible heat on precipitation may relate to its coupling with temperature and convective processes.
3. With rising temperatures in the future, the sensible heat will simultaneously increase and inhibit the formation of precipitation to some extent. In addition, the soil moisture has showed a decreasing trend, which would cause a reduction of precipitation in the dry periods. Under a changing climate, the impact of soil moisture on precipitation will further aggravate in the future.

The applicability of LSTNet in different climate regions (i.e., Central California) were investigated. The significance of validating LSTNet outputs with the GLACE-CMIP5 experiments in the future were also discussed. All in all, the LSTNet neural network can effectively capture the relative importance of key variables, such as sensible heat, in the LA interactions across different climate regions. The outputs are basically consistent with the physical process. It is crucial to assess the impact of sensible heat in various areas because it has the potential to inhibit precipitation formation and intensify the severity of drought conditions. Under climate change, the decline in soil moisture and the rise in sensible heat would further diminish precipitation, raise temperature and increase drought probability in the future. This study not only has enhanced our knowledge on the infleuntial mechanisms in the LA interactions during the drought evolution and intensification, but also provided valuable insights for further development and advancement of hydrologic models for drought monitoring and forecasting.

Chapter 6. Leveraging Atmospheric Conditions to Enhance the Drought Predictability over the Vietnamese Mekong Delta

Key Points:

- Atmospheric conditions in the external precipitation source region enhance the ability of deep neural networks to predict droughts
- The neural network effectively predicts meteorological and agricultural droughts, and compound dry-hot events at a 3-month lead time

Abstract

In the past few decades, severe drought events have profoundly affected the ecological, social, and economic aspects of the Vietnamese Mekong Delta (VMD). Given these impacts, accurate prediction of droughts in the VMD is essential to improve preparedness and optimize drought management and mitigation strategies. However, the potential benefits of leveraging atmospheric conditions from the external precipitation source region to enhance drought prediction accuracy are poorly understood. In this study, a Convolutional Gated Recurrent Unit (ConvGRU) neural network, was designed and utilized to evaluate whether the surrounding atmospheric conditions can enhance the performance of deep learning algorithms in predicting droughts over the VMD. The ConvGRU model incorporates the atmospheric conditions from the external precipitation source region during the model training and has demonstrated superior capability in predicting meteorological and agricultural droughts, as well as compound dry-hot events. Particularly at a 3-month lead time, it successfully predicts approximately 90% of meteorological drought events and about 80% of agricultural drought events, with fewer than 10% false predictions for drought months and events. Furthermore, ConvGRU predicts about 70% and 80% compound dry-hot months and events, respectively. ConvGRU effectively predicts the most severe meteorological drought and the longest agricultural drought at the lead time of 3 months but underestimates the severity of the most severe compound dry-hot event at the onset stage. The outstanding performance of the ConvGRU model in drought prediction at the 3-month lead time is likely due to the delayed influence of atmospheric conditions from the external precipitation source region, including specific humidity, U- and V-wind.

6.1 Introduction

As recurrent and widespread natural hazards, droughts exert significant impacts on water resource management, agricultural productivity, and socioeconomic stability across the globe (Mishra and Singh, 2010; Van Loon, 2015; Miralles et al., 2019). Records from the Emergency Events Database (EM-DAT) indicate that over 400 drought events have occurred worldwide in the 21st century, which affected more than 1.6 billion people and caused over 170 billion USD losses (Delforge et al., 2023). Specifically, the Mekong River Basin has experienced frequent and severe droughts over the past two decades, notably in 2002, 2005, 2010, 2015-2016, and 2019-2020, with significant socioeconomic consequences (Guo et al., 2017; Lee and Dang, 2018; Kang and Sridhar, 2021; Keovilignavong et al., 2021). In the Vietnamese Mekong Delta (VMD, as shown in Figure 6.1), suffered over 300 million USD in damages to agriculture and aquaculture during the 2015-2016 drought (Nguyen, 2017). The subsequent 2019-2020 drought caused significant water shortages and saltwater intrusion, impacting 82,000 households and exposing more vulnerable populations in VMD to significant water shortage risks (United Nations Resident Coordinator in Viet Nam, 2020). Under the context of climate change, the severity and frequency of droughts are projected to increase globally (Dai, 2013; Ault, 2020; Pokhrel et al., 2021; Seneviratne et al., 2021) as well as regionally (i.e., Mekong River Basin) (Li Y. et al., 2021; Dong et al., 2022; Kang et al., 2022). Therefore, accurate prediction of drought conditions in the VMD is crucial for strengthening drought preparedness, providing timely mitigation strategies, and enhancing drought resilience and adaptability.

To date, drought prediction methodologies can be categorized broadly into three main types: statistical methods, dynamical models, and hybrid approaches (Dikshit et al., 2021a; Nandgude et al., 2023). Statistical methods analyze the causal relationships between relevant variables and drought indices using historical data (Xu et al., 2018a; Barrett et al., 2020). Due to their simplicity and effectiveness, numerous statistical methods have been developed and utilized in drought prediction. For example, Yan et al. (2017) and Wu et al. (2022a, 2023) have advanced copula-based drought prediction methods that effectively predict seasonal hydrological and agricultural droughts. Furthermore, Zhang et al. (2021) demonstrated that incorporating ENSO into the meta-Gaussian model significantly enhances the predictability of agricultural drought in regions impacted by large-scale atmospheric circulations. Among statistical approaches, machine learning and deep learning models have gained prominence, including Random Forest (Park et al., 2019, Li J. et al., 2021a), Artificial Neural Networks (Le et al., 2017; Khan et al., 2020a, 2020b), Support Vector Machines (Belayneh et al., 2016; Tian et al., 2018; Khan et al., 2020a), Long Short-Term Memory networks (Poornima and Pushpalatha, 2019; Dikshit et al., 2021a, 2021b) and so on. Rhee and Im (2017) employed three machine learning models to predict meteorological drought for ungauged areas using long-range climate forecasts and remote sensing data. Additionally, Dikshit and Pradhan (2021) attempted to interpret the deep learning models in spatial drought prediction with the SHapley Additive exPlanations (SHAP). They suggested adding climatic variables as predictors in the prediction model because SHAP captures the importance of climate variables at different time scales, which align with the physical interpretations.

Dynamical prediction models, such as the National Centers for Environmental Prediction (NCEP) Coupled Forecast System model version 2 (CFSv2, Saha et al., 2014; Siegmund et al., 2015) and the European Centre for Medium-Range Weather Forecasts (ECMWF, Bonavita et al., 2016; Johnson et al., 2019), simulate real land-atmosphere-ocean interactions and processes based on climate and hydrologic models (Hao et al., 2018a; Xu et al., 2018a). In recent years, several multi-model ensembles have been developed and applied in precipitation or drought forecasting (Mo and Lyon, 2015; Ma et al., 2016; Xu et al., 2018b), such as the North American Multi-Model Ensemble (NMME, Kirtman et al., 2014). Generally, dynamical models were developed to simulate and forecast weather and climate conditions, with precipitation and temperature predictions being utilized to calculate drought indices (Dutra et al., 2014; Li et al., 2016; Hao et al., 2018a). Despite significant advances in short-term precipitation forecasting over the past few decades (Hapuarachchi et al., 2011; Ning et al., 2022), the monthly or seasonal accuracy of drought predictions remains constrained by the inherent unpredictability of the ocean-atmosphere system and gaps in physical understanding (Yoon et al., 2012; Hao et al., 2018a). As for the hybrid models, the critical idea is to integrate the strengths of both statistical and dynamical models in drought prediction (Xu et al., 2018a; Aghakouchak et al., 2022). Therefore, a pivotal aspect is the weighting assigned to the statistical and dynamical models. For example, Madadgar et al. (2016) integrated the NMME model with a Bayesian-based statistical approach using the Expert Advice algorithm (Cheng and AghaKouchak, 2015), achieving superior performance compared to the standalone NMME model.

Compared with dynamical models, statistical models, particularly deep learning models, have the advantages of efficiency and accuracy, which have been widely applied in drought prediction in recent years (Liu et al., 2024; Márquez-Grajales et al., 2024). The variables used in deep neural networks typically include hydrometeorological variables such as precipitation, temperature, and potential evapotranspiration, alongside climatic variables that characterize large-scale atmospheric-oceanic circulation patterns, such as sea surface temperature (SST) and Pacific Decadal Oscillation (PDO) (Zhang et al., 2019; Dikshit et al., 2022). Instead of simply utilizing time-series data of variables as model inputs, incorporating detailed spatial-temporal data may enhance the regional drought prediction capabilities of deep neural networks. For example, Holgate et al. (2020) pointed out that drought occurrence and intensification in southeast Australia were predominantly influenced by reduced oceanic moisture. Similarly, Stojanovic et al. (2021) and Zhou and Shi (2024) noted that precipitation moisture source for the VMD primarily originated from the China Seas during the dry season. This result means the atmospheric conditions from the external precipitation source region (e.g., China Seas) of the target area (i.e., VMD) would play a critical role in the occurrence and intensification of droughts. Thus, incorporating atmospheric conditions in the external precipitation source region into the prediction model, is likely to improve the performance of deep neural networks in predicting sub-seasonal to seasonal droughts. Therefore, this study aims to investigate the performance of deep learning algorithms in predicting drought, particularly meteorological and agricultural droughts, as well as compound dry-hot events, when considering the atmospheric conditions from the external precipitation source region.

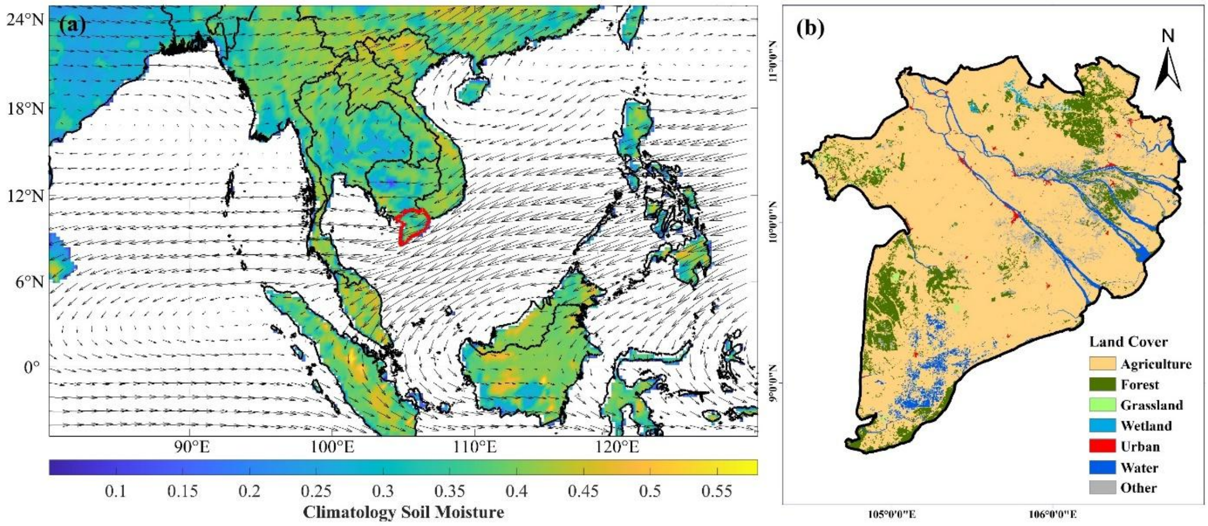

Figure 6.1. (a) Climatology soil moisture (shading) and moisture flux in the dry season (arrows) derived from ERA5 over the spatial domain of the precipitation source region of Vietnamese Mekong Delta (VMD, within the red line); (b) Geographical location and land cover types of VMD, based on the European Space Agency Climate Change Initiative land cover datasets.

6.2 Data

ERA5, the fifth-generation reanalysis data from the European Centre for Medium Range Weather Forecasts (ECMWF), covers the period from January 1940 to the present (Hersbach et al., 2023a, 2023b). The ERA5 data used in this study include both the variables (i.e., specific humidity, temperature, geopotential, U- and V-component of wind) at four pressure levels (i.e., 200, 600, 850, and 1000 hPa) and variables at the land surface (i.e., total precipitation, surface sensible heat flux, surface latent heat flux, volumetric soil water layer 1 and 2, temperature at 2 meters, U- and V-component of wind at 10 meters, and sea surface temperature). Here, two layers of volumetric soil water were aggregated into one layer. Therefore, a total of 28 input variables were used in this study (Table 6.1). All variables were processed at a monthly interval with a spatial resolution of 0.25° by 0.25°. The specific temporal span of ERA5 data employed in this analysis ranges from January 1940 to November 2023. Approximately 70% of the dataset (from January 1940 to December 1999) was used for model training, while 20% (from January 2000 to December 2015) was applied for validation, and the remaining 10% (from January 2016 to November 2023) was for testing. The ERA5 reanalysis data used in this study are freely available from the Copernicus Climate Change Service (C3S) Climate Data Store (https://cds.climate.copernicus.eu/cdsapp#!/home).

Table 6.1. The 28 input variables derived from ERA5 data for model training.

Data	Types	Variables	Inputs	Time span
ERA5	Pressure levels	Specific humidity Temperature Geopotential U component of wind V component of wind	Monthly interval 4 pressure levels (200, 600, 850 and 1000 hPa)	January 1940 - November 2023
	Single level	Total precipitation Surface sensible heat flux Surface latent heat flux Temperature at 2 meters U-wind at 10 meters V-wind at 10 meters Sea surface temperature	Monthly interval	
		Volumetric soil water layer 1 Volumetric soil water layer 2	Monthly interval (aggregate into 1 layer)	

6.3 Methods

6.3.1 Deep neural networks

In this study, the Convolutional Gated Recurrent Unit (ConvGRU) deep learning model was proposed to predict drought conditions over the VMD. The neural network architecture, illustrated in Figure 6.2, begins with the convolutional neural network (CNN) layers (LeCun et al., 1998; Krizhevsky et al., 2012), augmented by Spatial-Channel wise attention mechanisms (Woo et al., 2018). The input data shape contains 18-time steps, each involving 28 variables across the spatial domain of 120 by 200 grid cells (25°N-5°S, 80°E-130°E, covers VMD and its surrounding areas as shown in Figure 6.1(a)). The CNN layers extract the dependencies between variables (Conv3d layer) and spatial information (Conv2d and Spatial-Channel attention layers) that may contribute to the prediction of variables over the VMD. The afterwards block is a Gated Recurrent Unit layer (GRU, Chung et al., 2014), which extracts temporal variation from previous time steps. The final component of the network, the Multilayer Perceptron layer (MLP, Satori and Antsaklis, 1991), outputs the multivariate predictions. To evaluate if the data from the external area (outside of the VMD in Figure 6.1(a)) can influence the prediction of variables in the VMD, two scenarios, ConvGRU_FULL and ConvGRU_VMD, were trained based on this neural network structure. The difference between these two scenarios is that the inputs of ConvGRU_FULL covers all data and information from the VMD and the external area, while the inputs of ConvGRU_VMD only preserves data within the VMD, with data from the external area specified as 0. In this study, the spatial average of variables (excluding SST) over the VMD was taken as the prediction target, and the 27 variables were predicted at lead times of 1, 3, 6, and 12 months.

To further explore the contribution of different variables in drought predictions based on the neural networks, a simplified approach was used to calculate the relative importance of each variable in the ConvGRU_FULL scenario. The difference between ConvGRU(X_0) and ConvGRU(X_i) was used to represent the weight of each variable in the model, where $X_0 = [1, 1, 1, \ldots, 1]$, X_i is similar to X_0 but the i^{th} variable set to 0, e.g., $X_2 = [1, 0, 1, \ldots, 1]$. The length of X_0 and X_i corresponds to the number of variables in the input data. Furthermore, data in grid cells from local (i.e., VMD) and external (outside of the VMD) areas were prescribed as 0, respectively, to assess the relative importance of information from local and external areas.

6.3.2 Standardized drought index

Over the past decades, a variety of drought indices have been developed to quantify drought events, each with strengths and weaknesses (Mishra and Singh, 2010; Zhou et al., 2021). In this study, standardized drought indices were used to evaluate the performance and accuracy of deep neural network predictions. Specifically, the Standardized Precipitation Index (SPI, McKee et al., 1993; Edwards and McKee, 1997) and the Standardized Soil Moisture Index (SSI, Hao and AghaKouchak, 2013; Javed et al., 2021) were utilized to assess meteorological and agricultural droughts, respectively. Subsequently, the Standardized Precipitation Temperature Index (SPTI) was derived by coupling the SPI with the Standardized Temperature Index (STI, Zscheischler et al., 2014; Li J. et al., 2021b) through the copula function (i.e., t-copula) to evaluate the conditions associated with compound meteorological droughts and heatwaves (Hao et al., 2018b; Li J. et al., 2021b). The SPI, SSI, and SPTI were calculated on a monthly scale in this study.

6.3.3 Evaluation metrics

In this paper, the Pearson Correlation Coefficient (PCC), Root Mean Square Error (RMSE), Probability of Detection (POD), False Alarm Rate (FAR), and Absolute Bias (ABias) were used to measure the models' performance, as illustrated in Table 6.2.

The POD and FAR metrics were applied to assess the models' effectiveness at predicting drought on both monthly and event-based scales. The ABias evaluates whether the model can accurately capture the characteristics of drought events (i.e., onset, duration, intensity, and severity). The definition of drought events and drought characteristics are based on drought indices (as shown in Figure 6.3). Drought severity was defined as the cumulative sum of drought index values below the threshold for each event (i.e., the area in orange in Figure 6.3). Only drought events with a severity lower than -2 were taken into account in this study, for instance, the drought event with yellow color in Figure 6.3 was excluded. If a drought month or event occurs one month after (before) a drought event, it is considered as the same drought event (e.g., Drought 1 in Figure 6.3). Drought duration was defined as the number of consecutive months associated with each drought event, and drought intensity was identified as the value of the drought index lower than the threshold in the most severe drought month in each drought event (Figure 6.3).

Table 6.2. Equations and definitions for PCC, RMSE, POD, FAR, and ABias.

Metrics	Formula	Definitions
PCC	$\frac{\sum(O_i-\overline{O})(S_i-\overline{S})}{\sqrt{\sum(O_i-\overline{O})^2}\sqrt{\sum(S_i-\overline{S})^2}}$	O_i and S_i represent reference and prediction data
RMSE	$\sqrt{\frac{1}{n}\sum(O_i-S_i)^2}$	
POD	$\frac{C}{C+I}$	C and I represent the number of hit and miss drought months or events
FAR	$\frac{F}{F+I}$	F and I represent the number of false alarmed and hit no drought months or event
ABias	$\frac{\sum abs(O_i-S_i)}{N}$	O_i and S_i represent reference and predicted drought characteristics, N represents the number of actual drought events.

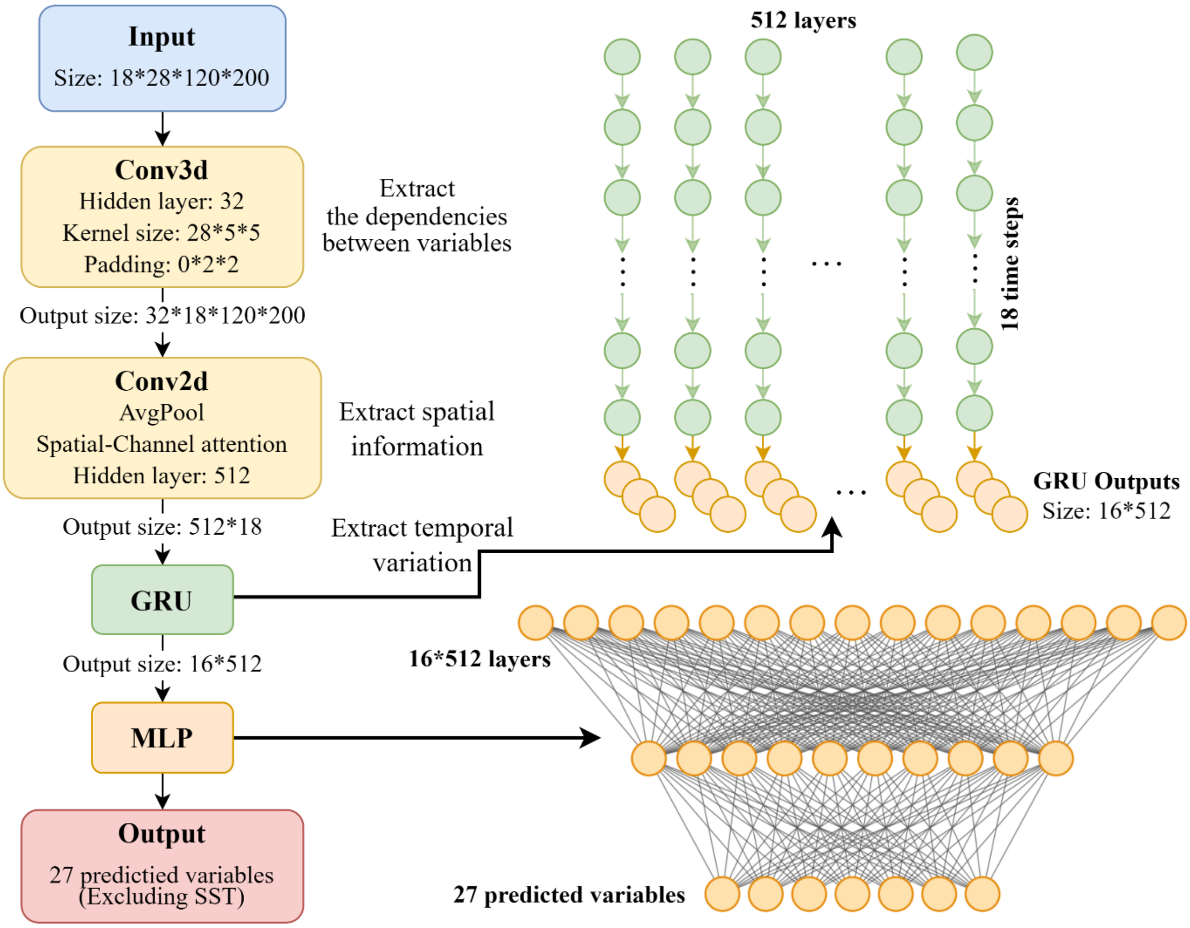

Figure 6.2. The structure of the ConvGRU neural network in this study, with CNN layers extracting the dependencies between variables and spatial information and GRU layer extracting temporal variation.

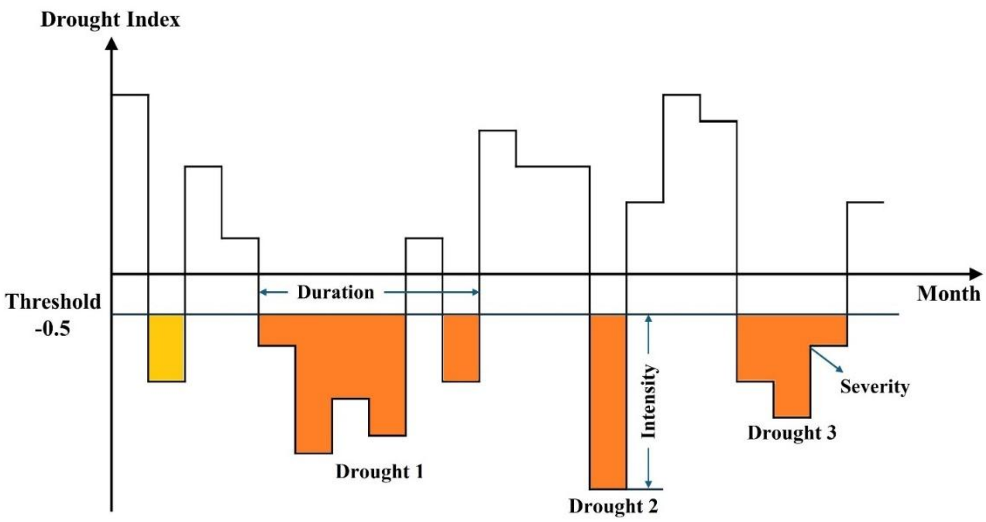

Figure 6.3. The definition of drought events and characteristics in this study.

6.4 Results

6.4.1 The performance of ConvGRU in predicting meteorological droughts

The ConvGRU_FULL demonstrates superior performance in predicting precipitation compared to the ConvGRU_VMD in terms of the metrics of PCC and RMSE across the four lead times (1, 3, 6, and 12 months, Table 6.3). The PCC values for ConvGRU_FULL range from 0.89 to 0.97 and reach the maximum on the Train set at 3 months lead time. Meanwhile, the ConvGRU_VMD's PCC values are predominantly below 0.93. For the RMSE metric, ConvGRU_FULL's predictions vary from 0.96 to 1.77 mm/day, while ConvGRU_VMD's RMSE values are higher, ranging from 1.46 to 2.03 mm/day across the four lead times. Notably, both scenarios exhibit their best performance at the lead time of 3 months rather than 1 month. The underlying reasons for this pattern are further explored in Section 6.5.1.

As illustrated in Figure 6.4, ConvGRU_FULL outperforms another scenario across most metrics at different lead times, especially at the lead time of 3 months. Consistent with the results in Table 6.3, ConvGRU_FULL achieves its optimal performance in predicting meteorological drought at the lead time of 3 months. It successfully identifies over 60% of drought months (SPI < -0.5), roughly 70% of severe drought months (SPI < -1.5), and about 90% of meteorological drought events, with fewer than 10% false predictions for both drought months and events at a 3-month lead time. As for the drought characteristics, at a 3-month lead time, ConvGRU_FULL well predicted the onset of meteorological drought events with the ABias of only 0.65 months. The ABias values for drought duration, intensity and severity are approximately 2 months, 0.5 and 1.4, respectively. In contrast, while the ConvGRU_VMD model is somewhat less proficient at detecting drought events and their onset, it accurately predicts the severity of droughts for the events it does identify, particularly at 6 and 12 months lead times, achieving an ABias of around 1.35.

6.4.2 The performance of ConvGRU in predicting agricultural droughts

In assessing the accuracy of predicting agricultural drought, the two scenarios, ConvGRU_FULL and ConvGRU_VMD, demonstrate performances similar to those observed for meteorological drought, as detailed in Table 6.4 and Figure 6.5. At the lead time of 1 month, both scenarios exhibit comparable results in terms of PCC and RMSE. However, ConvGRU_FULL outperforms ConvGRU_VMD at longer lead times of 3, 6, and 12 months, with PCC values exceeding 0.95 and RMSE values ranging from 0.008 to 0.2 m^3/m^3. Notably, ConvGRU_FULL shows distinct advantages, particularly at the 3-month lead time, where it successfully predicts approximately 80% of agricultural drought months and events, and 70% of severe drought months. ConvGRU_FULL excels not only in POD but also in FAR, with only about 6% of drought months being falsely predicted at the lead time of 3 months. When examining the characteristics of agricultural droughts, ConvGRU_FULL's superior performance continues at a 3-month lead time, with an ABias of 0.75 months for onset, 1.56 months for duration, 0.62 for intensity, and 1.06 for severity. Conversely, the ConvGRU_VMD model exhibits slightly enhanced performance compared to ConvGRU_FULL at the 1, 6, and 12-month lead times. For example, at a 1-month lead time, the ABias of the ConvGRU_VMD model in predicting agricultural drought intensity is less than 0.5, indicating an advantage in shorter-term predictions.

Table 6.3. The performance of ConvGRU in predicting precipitation at four lead times.

Metric	Lead times	1 month		3 months		6 months		12 months	
	Models	ConvGRU_FULL	ConvGRU_VMD	ConvGRU_FULL	ConvGRU_VMD	ConvGRU_FULL	ConvGRU_VMD	ConvGRU_FULL	ConvGRU_VMD
PCC	Train	0.91	0.90	0.97	0.93	0.92	0.91	0.91	0.91
	Valid	0.91	0.89	0.92	0.91	0.91	0.88	0.91	0.89
	Test	0.89	0.89	0.93	0.90	0.91	0.88	0.92	0.91
RMSE/ mm/day	Train	1.60	1.82	0.96	1.46	1.56	1.66	1.71	1.66
	Valid	1.62	1.78	1.49	1.63	1.59	1.84	1.57	1.77
	Test	1.77	1.84	1.51	1.76	1.62	2.03	1.55	1.64

Table 6.4. The performance of ConvGRU in predicting soil moisture at four lead times.

Metric	Lead times	1 month		3 months		6 months		12 months	
	Models	ConvGRU_FULL	ConvGRU_VMD	ConvGRU_FULL	ConvGRU_VMD	ConvGRU_FULL	ConvGRU_VMD	ConvGRU_FULL	ConvGRU_VMD
PCC	Train	0.96	0.96	0.99	0.97	0.97	0.95	0.96	0.95
	Valid	0.97	0.95	0.97	0.95	0.96	0.92	0.96	0.93
	Test	0.97	0.96	0.97	0.93	0.95	0.92	0.97	0.95
RMSE/ m^3/m^3	Train	0.019	0.019	0.008	0.015	0.016	0.018	0.016	0.020
	Valid	0.018	0.020	0.016	0.019	0.017	0.022	0.016	0.022
	Test	0.014	0.018	0.016	0.023	0.020	0.024	0.013	0.019

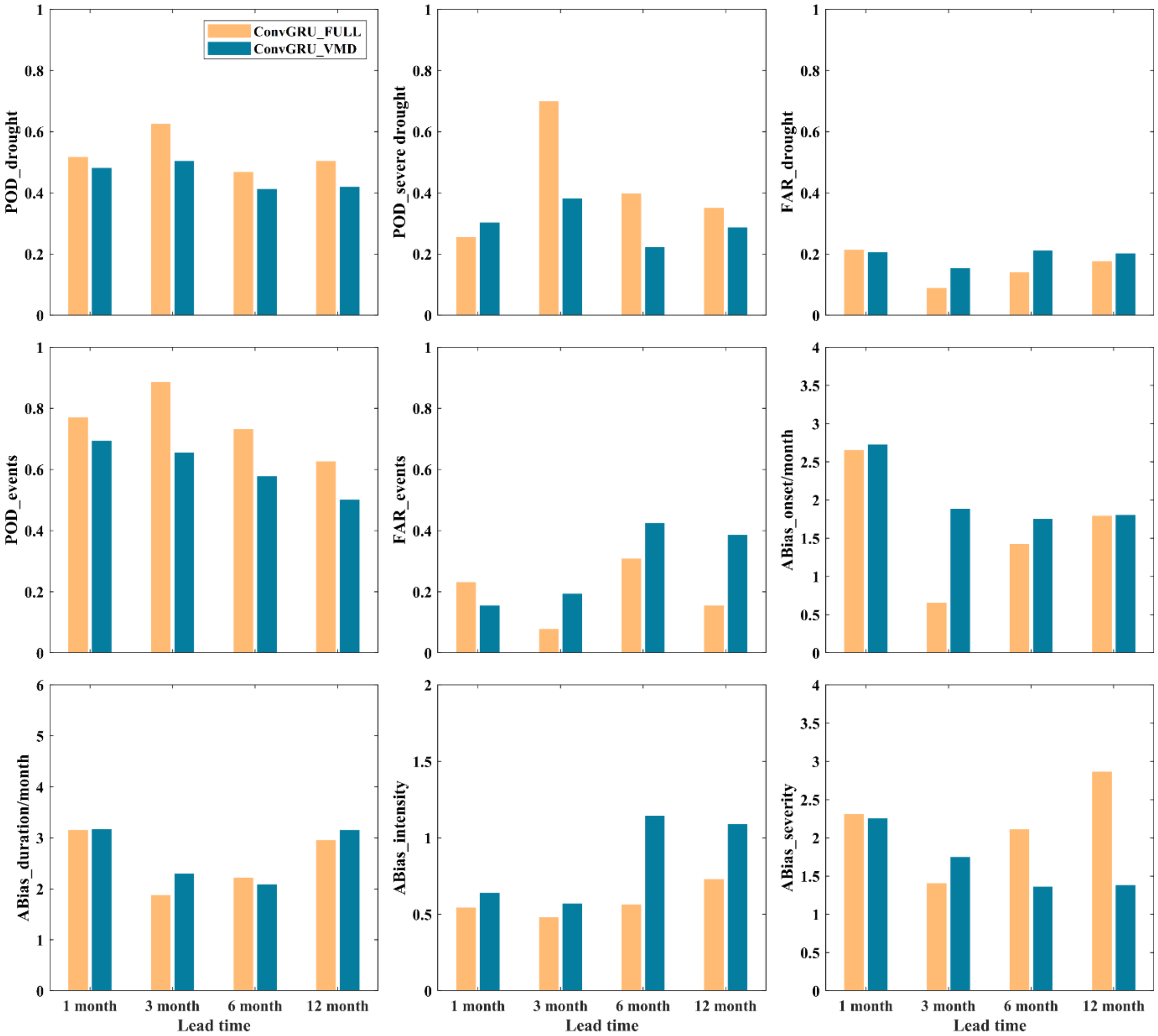

Figure 6.4. The performance of ConvGRU in predicting meteorological drought at four lead times (The definition and calculation of metrics refer to Section 6.3.3).

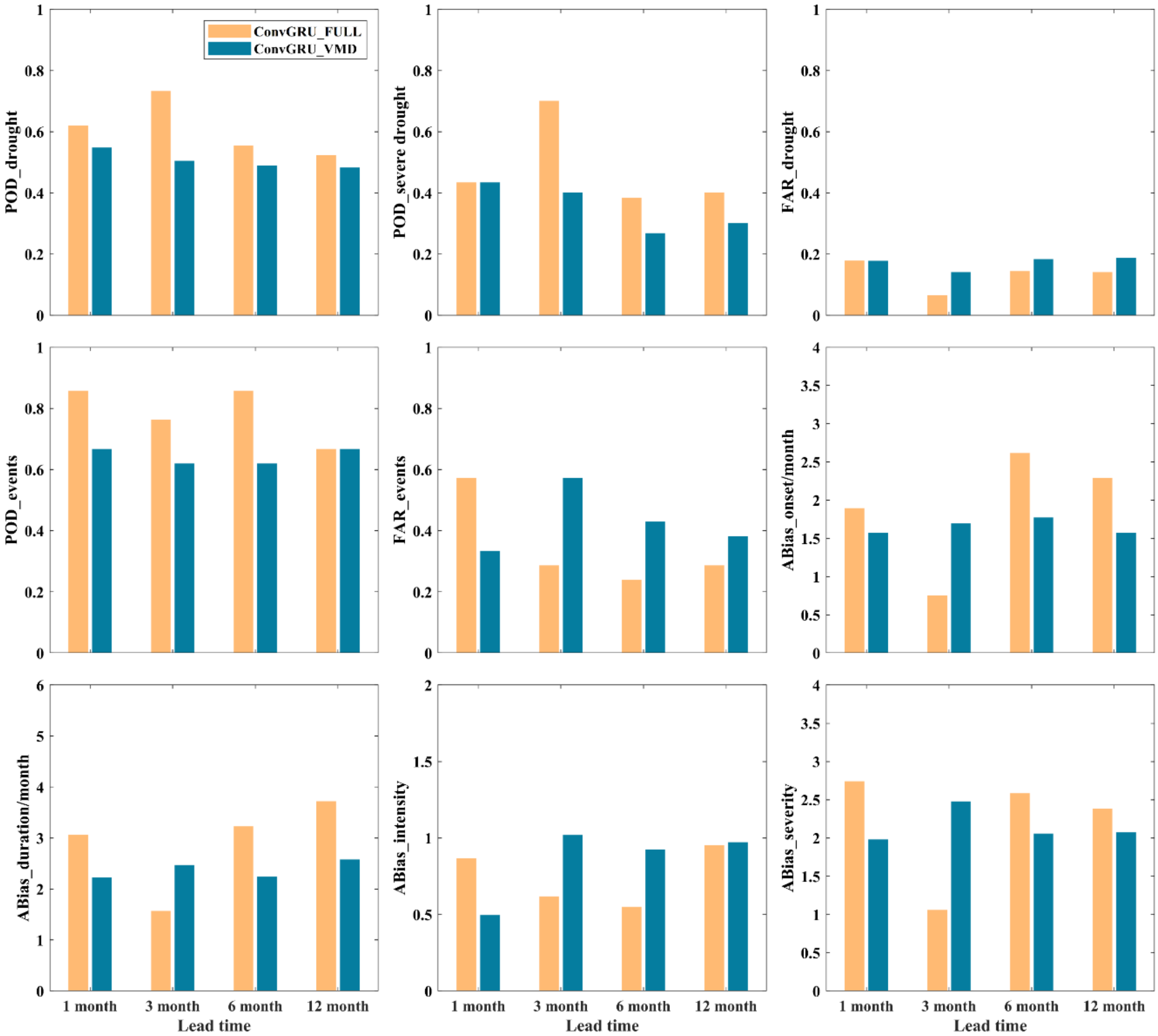

Figure 6.5. The performance of ConvGRU in predicting agricultural drought at four lead times.

6.4.3 The performance of ConvGRU in predicting compound dry-hot events

With the intensification of climate change, compound dry-hot events are increasingly recognized as a prevalent drought-related hazard worldwide (Diffenbaugh et al., 2017; Feng et al., 2020). Therefore, this study assesses the capability of ConvGRU to predict these complex events. Table 6.5 presents the metrics used to evaluate the performance of the model in predicting surface air temperature (i.e., temperature at 2 meters). ConvGRU_FULL consistently outperforms ConvGRU_VMD in temperature prediction, particularly at longer lead times (3, 6, and 12 months). Similar to the results in Section 6.4.1, ConvGRU_FULL achieves the best performance on the Train set at a 3-month lead time (PCC = 0.98, RMSE = 0.26 ℃).

As depicted in Figure 6.6, ConvGRU_FULL generally outperforms the ConvGRU_VMD model in capturing compound dry-hot months and events. Specifically, ConvGRU_FULL captures 67% compound dry-hot months (SPTI < -0.5), 55% severe compound dry-hot months (SPTI < -1.5) and falsely predicts 15% compound dry-hot months at 3 months lead time, marking the best performance across both models and all four lead times. ConvGRU_FULL excels at the 6-month lead time in terms of event-based POD and FAR (i.e., POD_events = 0.83, and FAR_events = 0.22). Regarding the characteristics of compound dry-hot events, ConvGRU_FULL has better performance in capturing the onset of compound dry-hot events with the ABias of about 2 months. The models' performance in predicting the duration of compound dry-hot events varies depending on the lead times. For example, ConvGRU_VMD excels at 3- and 12-month lead times with a similar ABias of around 3 months, while the ConvGRU_FULL performs better at 1- and 6-month lead times, with ABias values of 3.5 and 4.9 months, respectively. Generally, ConvGRU_FULL and ConvGRU_VMD effectively capture the intensity and severity of compound dry-hot events with the ABias below 0.5 and 3.5, respectively. Overall, the ConvGRU neural network exhibits slightly less accuracy in predicting compound dry-hot events compared to meteorological and agricultural droughts. This is largely attributed to the use of the copula-based SPTI, which inherently couples the errors and uncertainties from the predicted SPI and STI.

Table 6.5. The performance of ConvGRU in predicting surface air temperature at four lead times.

Metric	Lead times	1 month		3 months		6 months		12 months	
	Models	ConvGRU_FULL	ConvGRU_VMD	ConvGRU_FULL	ConvGRU_VMD	ConvGRU_FULL	ConvGRU_VMD	ConvGRU_FULL	ConvGRU_VMD
PCC	Train	0.91	0.92	0.98	0.94	0.93	0.92	0.91	0.91
	Valid	0.90	0.87	0.91	0.92	0.90	0.87	0.89	0.87
	Test	0.90	0.83	0.89	0.84	0.87	0.88	0.92	0.83
RMSE/ °C	Train	0.74	0.49	0.26	0.37	0.48	0.42	0.45	0.48
	Valid	0.54	0.49	0.40	0.38	0.41	0.62	0.44	0.49
	Test	0.54	0.60	0.47	0.55	0.49	0.91	0.36	0.52

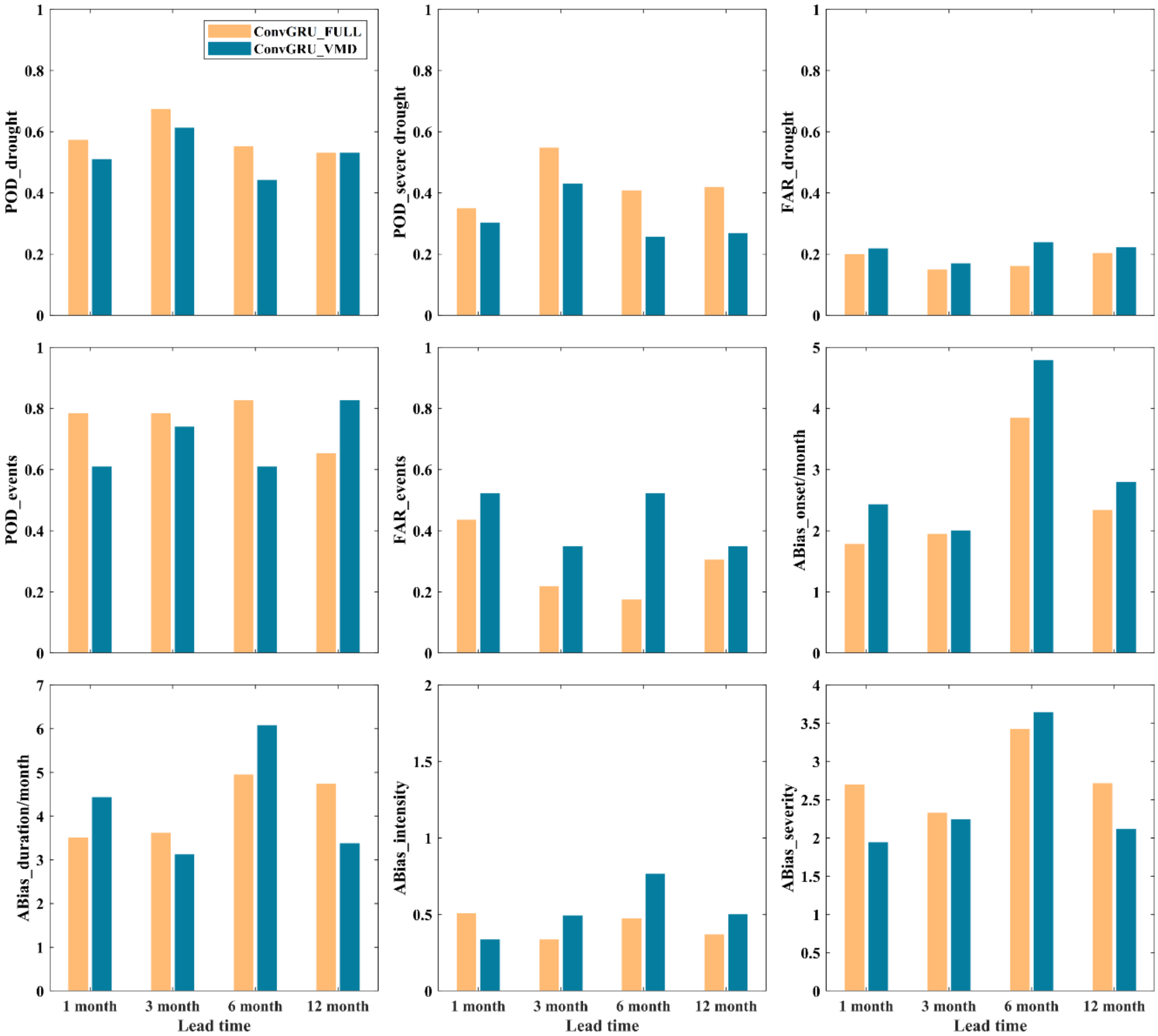

Figure 6.6. The performance of ConvGRU in predicting compound dry-hot events at the four lead times.

6.4.4 The performance of ConvGRU in predicting historic events

In addition to evaluating the performance of ConvGRU through metrics (e.g., PCC, POD) in previous sections, three representative events are selected in this section to explore the capabilities of deep neural networks in predicting historic events. As seen in Figure 6.7, the first event is the most severe meteorological drought event, spanned from May 1950 to October 1951, lasting 18 months with a drought severity of -17.86. The ConvGRU_FULL accurately predicts the onset and end of this drought at a lead time of 3 months. Additionally, the SPI predicted by the ConvGRU_FULL model aligns with the observation from the ERA5 data, and the absolute severity bias of the ConvGRU_FULL's prediction for this drought event is only 0.89. In the meantime, the ConvGRU_VMD performs slightly worse than ConvGRU_FULL with an absolute severity bias of 1.71. The second event in Figure 6.7 is the longest agricultural drought event, starting in October 1957 and ending in April 1959. The duration and severity of this event are 20 months and -10.62, respectively. Similarly, the ConvGRU_FULL model accurately predicts the onset and end of this drought, while the ConvGRU_VMD model captures only the mid-term conditions of this agricultural drought event. The absolute severity biases for these two scenarios regarding this event are 1.43 and 6.7, with absolute intensity biases of 0.11 and 0.66, respectively. The third representative event, the most severe compound dry-hot event, began in March 2015 and ended in January 2017, lasting 23 months. The ConvGRU_FULL captures the initiation of this event, but the severity of the onset stage (March to August 2015) of the event is not well predicted by ConvGRU_FULL. The main reason for this problem is that the ConvGRU_FULL overestimates precipitation (1.07 mm/day) and underestimates temperature (-0.42 ℃) in this period, causing an underestimated severity of this event. The total absolute severity bias for ConvGRU_FULL in predicting this event is 7.23, mainly due to the bias in the onset stage, which is 5.59. ConvGRU_VMD performs even worse than ConvGRU_FULL with the bias of 8.54. Overall, ConvGRU_FULL accurately predicted the most severe meteorological drought and the longest agricultural drought at the lead time of 3 months, and its performance is better than the ConvGRU_VMD. As for the prediction of the most severe compound dry-hot event, the ConvGRU_FULL performs not as well as in predicting the meteorological and agricultural drought events but is still better than the ConvGRU_VMD model.

In addition to these three representative events, we further assessed the ConvGRU_FULL's ability to reflect historical severity variations in meteorological drought and compound dry-hot events (Figure 6.8). The severity of meteorological drought predicted by the ConvGRU_FULL generally follows the observed variations based on the ERA5 data. Before 2000, the ConvGRU_FULL slightly overestimates the meteorological drought severity with the precipitation bias of 0.25 mm/day (averaged bias before 2000). On the contrary, it slightly underestimates the severity after 2000 with a bias of -0.02 mm/day. Similarly, a clear transition point occurs around 2000 in the severity prediction of compound dry-hot events. The ConvGRU_FULL overestimates the severity of compound dry-hot events before 2000 with a temperature bias of 0.11 ℃ (averaged bias before 2000) and underestimates the severity after 2000 with a bias of -0.17 ℃. This phenomenon primarily results from the use of training data spanning from 1940 to 1999, which does not adequately capture the further climate change developments in the 21st century.

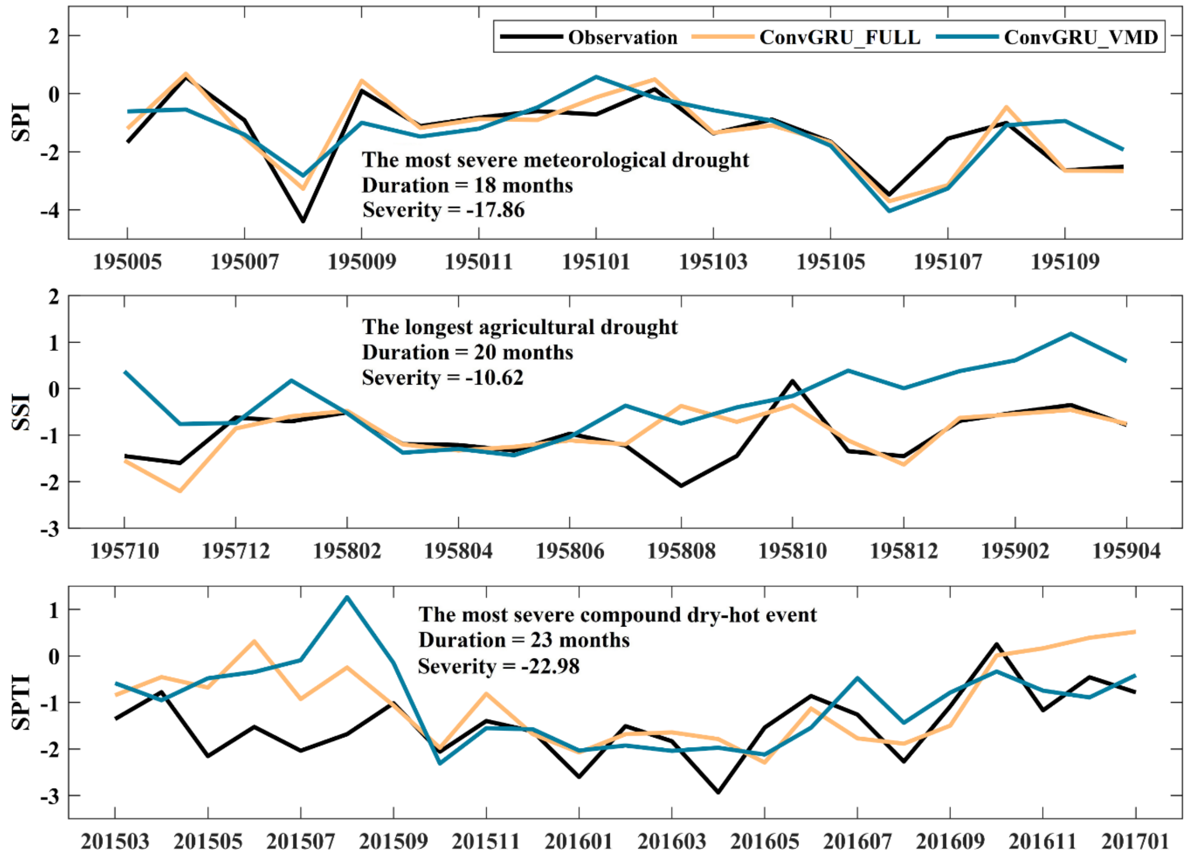

Figure 6.7. The performance of ConvGRU in predicting three historic events at a 3-month lead time. Namely the most severe meteorological drought event (top), the longest agricultural drought event (middle) and the most severe compound dry-hot event (bottom).

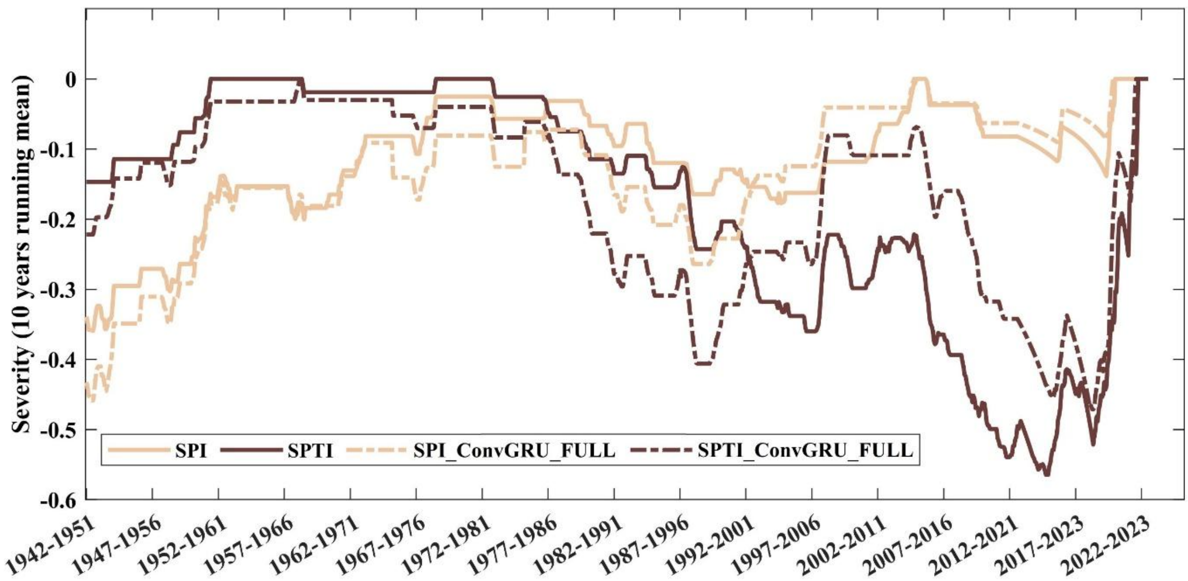

Figure 6.8. Historical and ConvGRU_FULL predicted (3-month lead) variations in the severity of meteorological drought and compound dry-hot events.

6.5. Discussion

6.5.1 Why does ConvGRU perform best at the lead time of 3 months?

Previous studies have demonstrated that deep neural networks generally achieve optimal performance at short lead times, typically at a 1-time step, with their accuracy progressively declining as the lead time extends (Dikshit et al., 2021b; Wu et al., 2022b). For instance, the deep neural network developed by Khan and Maity (2024) exhibits its best performance at a 1-month lead time for hydrological drought prediction. However, according to the results in Section 6.4, the ConvGRU_FULL model typically performs best at the lead time of 3 months rather than 1 month across the metrics. Therefore, to explore the reasons behind the superior performance of ConvGRU_FULL at this specific lead time, the relative importance of variables was calculated in both local and external areas for the prediction of precipitation based on the method described in Section 6.3.1. Figure 6.9 illustrates the ratio of the relative importance of variables from external area to local area. The higher the ratio, the more important the variables from the external precipitation source region. Notably, for specific humidity and V-wind, the ratio increases over the extension of lead time, which reveals that specific humidity and V-wind from the external area become more influential in predictions over the 3-, 6- and 12-month lead times compared with the 1-month lead time. In addition, ConvGRU_FULL emphasizes the significance of U-wind from the external area in predicting precipitation at the 3-month lead time. Here, specific humidity, U-wind and V-wind are variables controlling the moisture transport process (van der Ent et al., 2013; Guan et al., 2022), ultimately affecting the amount of precipitation in the VMD (Stojanovic et al., 2021; Zhuo and Shi, 2024). Therefore, the enhanced performance of the ConvGRU_FULL model in drought prediction at a 3-month lead time is likely attributable to the delayed effects of atmospheric conditions (i.e., specific humidity, U-wind, V-wind) in the external area on the VMD drought conditions through the process of moisture transport. Conversely, temperature and geopotential from the external area do not exhibit delayed effects on the prediction of VMD precipitation. Although the relative importance of specific humidity and V-wind from the external area is more significant at 6- and 12-month lead than 3-month lead time, the performance of neural networks gradually deteriorates with the extension of lead time. Consequently, the ConvGRU_FULL model achieves its best performance at the 3-month lead time.

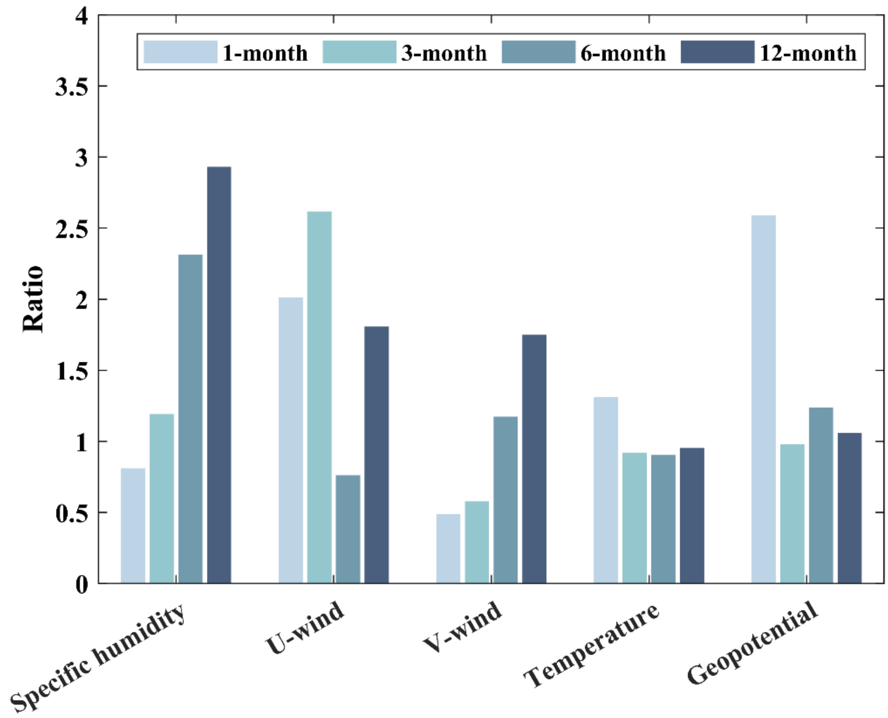

Figure 6.9. The ratio of the relative importance of variables from external area to local area in predicting precipitation at 1-, 3-, 6- and 12-month lead times in the ConvGRU_FULL model.

6.5.2 The effect of data splitting on model training and drought prediction

In Figure 6.8 in Section 6.4.4, a clear transition point was identified around 2000 in predicting the severity of meteorological droughts and compound dry-hot events. The ConvGRU_FULL model overestimates the severity before 2000 and underestimates the severity afterwards. This issue arises possibly because the data for model training only covers 1940 to 1999, which fails to effectively capture the further developments of climate change in the 21st century. In previous studies, k-fold cross-validation is generally used to avoid the appearance of spurious effects of any particular partition of the input data and to improve the model's generalization ability (Markatou et al., 2005; Javier and Manuel, 2020). However, the substantial input data in this study makes k-fold cross-validation time-consuming. Therefore, this study proposes to train the ConvGRU_Random model by randomly selecting 70% of all data (1940-2023) as the training input for each epoch. As shown in Figure 6.10, the severity of meteorological drought predicted by the ConvGRU_Random model generally follows the observation based on the ERA5 data. Compared with the ConvGRU_FULL model, ConvGRU_Random performs better in predicting compound dry-hot events, especially from 1950 to 2000, aligning closely with the ERA5-based observations. There is no significant transition point in the severity prediction of both meteorological drought and compound dry-hot events. This suggests that randomly selected input data may enhance the model performance and help to capture climate change throughout the entire period. However, there are still several questions that are worth discussing and exploring in the future. First, will randomly selecting input training data for each epoch introduce uncertainty, and would using the same randomly selected data for each epoch increase stability? Second, as seen in this study, patterns from nearly a century ago may not accurately reflect today's relationships among variables. In addition, Figure 6.11 represents the variations of annual uncertainty of precipitation, soil moisture, and sensible and latent heat in the ERA5 dataset from 1940 to 2020, indicating a significant reduction in the uncertainty and improvement in data quality from 1940 to 1979. Similarly, this trend continues, as described by Hersbach et al. (2020), with the uncertainty in temperature, zonal wind, and specific humidity showing sustained declines from 1979 to 2018. Therefore, considering the data quality, whether extensive long-term historical data is beneficial or detrimental to the development of deep learning models for drought prediction still needs to be investigated and explored.

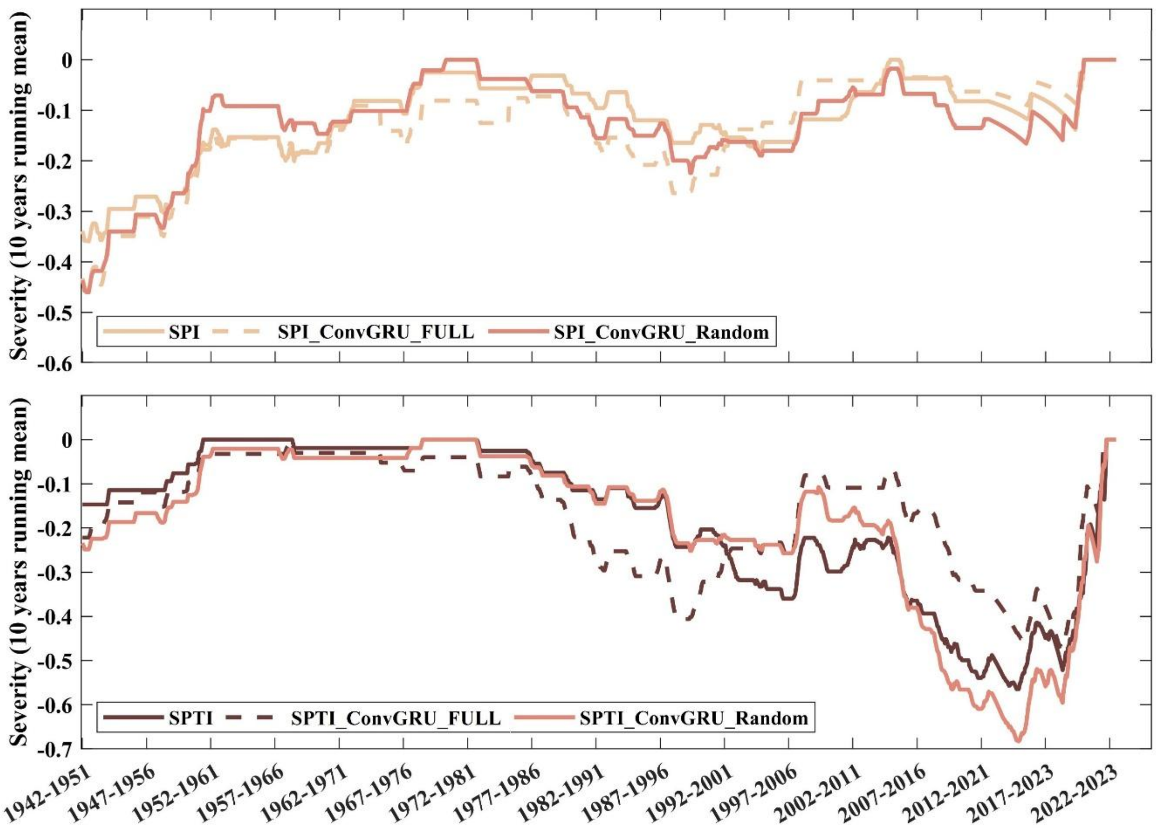

Figure 6.10. Historical and ConvGRU predicted (3 months lead) variations in the severity of meteorological drought and compound dry-hot events. ConvGRU_Random means training the ConvGRU model with randomly selected 70% of the data as input.

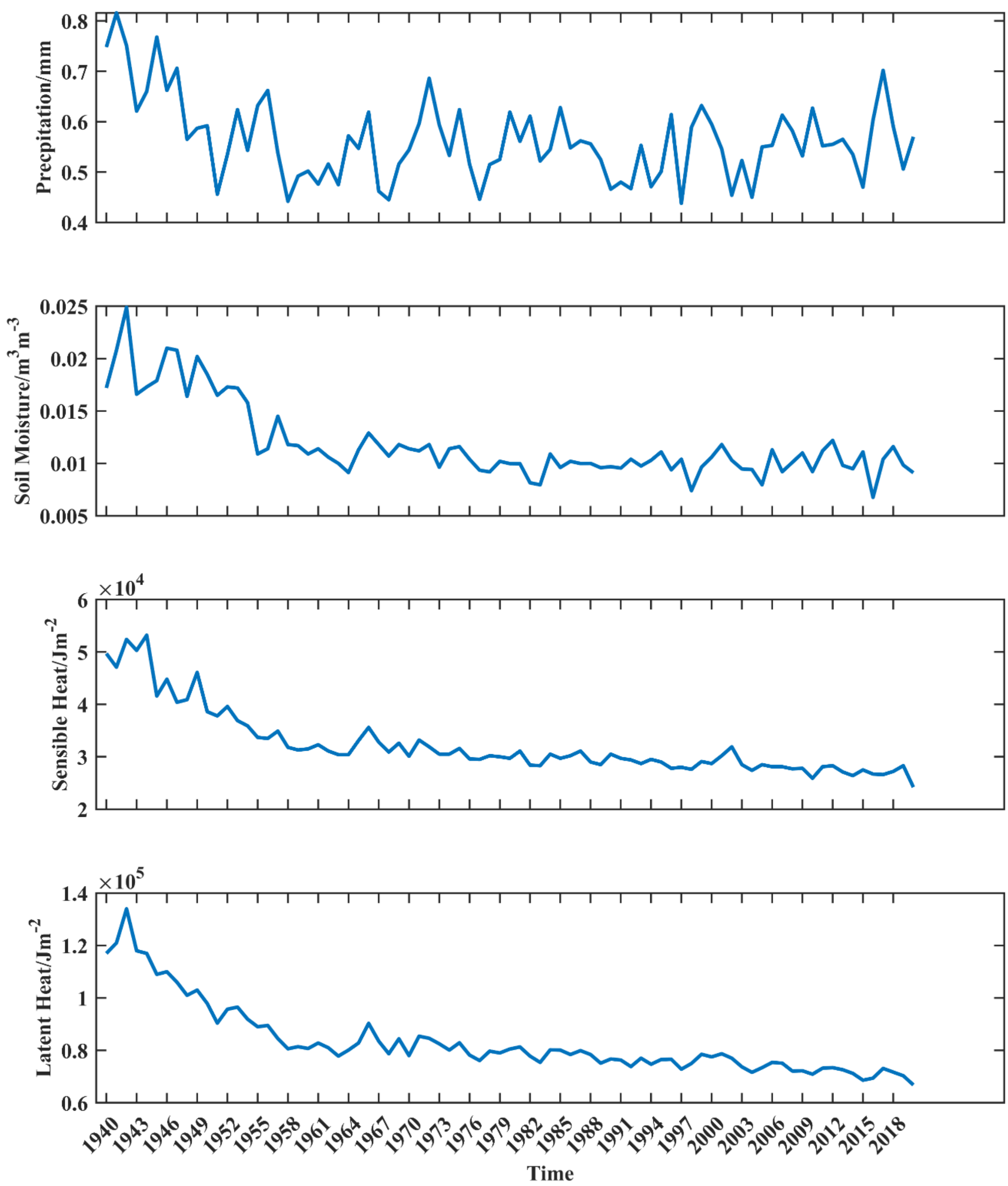

Figure 6.11. Annual variations in the uncertainty of precipitation, soil moisture, and sensible and latent heat from 1940 to 2020.

6.6 Conclusion

In this study, ConvGRU neural network was designed and evaluated to explore whether the atmospheric conditions from the external precipitation source region could enhance the performance of deep learning algorithms in predicting droughts over the VMD. The primary findings are summarized below:

1. ConvGRU_FULL outperforms the ConvGRU_VMD model in predicting meteorological and agricultural droughts across the four lead times, capturing about 90% of meteorological drought events and about 80% of agricultural drought events at a 3-month lead time. ConvGRU_FULL excels not only in POD but also in FAR, with fewer than 10% drought months and events being falsely predicted at the lead time of 3 months. In terms of the prediction of compound dry-hot events, ConvGRU_FULL has better performance in capturing compound dry-hot months and events, captured 67% compound dry-hot months and 55% severe compound dry-hot months, and about 80% compound dry-hot events at 3 months lead time.

2. At the 3-month lead time, ConvGRU_FULL accurately predicted the onset of meteorological and agricultural drought events with the ABias of approximately 0.7 months. The ABias for drought duration is below 2 months, and the ABias for intensity and severity are lower than 0.62 and 1.4, respectively. The ConvGRU neural network demonstrates slightly lower accuracy in predicting compound dry-hot events compared to meteorological and agricultural droughts, which is primarily due to the copula-based SPTI integrating the errors and uncertainties from the predicted SPI and STI.

3. ConvGRU_FULL effectively predicted the most severe meteorological drought and the longest agricultural drought at the lead time of 3 months but underestimates the severity of the onset stage of the most severe compound dry-hot event.

4. The superior performance of the ConvGRU_FULL model in drought prediction at the 3-month lead time is likely attributable to the delayed effects of atmospheric conditions (i.e., specific humidity, U-wind, V-wind) in the external area on the VMD drought conditions through the process of moisture transport.

Overall, incorporating the atmospheric conditions from the external precipitation source region significantly improves the ConvGRU neural network's capability to predict droughts in the VMD droughts, particularly at the lead time of 3 months. These findings underscore the critical role of external atmospheric conditions in improving the accuracy of sub-seasonal to seasonal drought predictions. This enhancement is vital for the development of an effective early warning system for droughts, providing a crucial tool for anticipatory actions and management strategies.

Chapter 7. Discussion and Conclusion

This chapter synthesizes the significant research contributions of this thesis. It methodically addresses each research question delineated in Chapter 1, identifies the limitations of this research, and proposes directions for future research grounded in the findings presented. Finally, the chapter concludes with a summary of the primary insights derived from this study.

7.1 Contributions of the research

Based on the detailed analyses and investigations conducted in Chapters 4 to 6, the research objectives and questions outlined in Chapter 1 have been comprehensively addressed. The primary contributions are listed point by point as follows:

Objective 1: To elucidate the source regions of precipitation moisture and identify the dominant factors influencing these processes during drought periods in the VMD, as detailed in Chapter 4. Research questions include:

- What are the characteristics of the moisture source regions affecting the VMD precipitation?

Atmospheric moisture transport plays a crucial role in the dynamics and onset of droughts. Therefore, understanding the origins of precipitation moisture is essential for enhancing predictions and comprehension of drought initiation in the VMD.

The analysis in Chapter 4 reveals that contribution evaporation from external regions substantially surpasses that from local sources. From 1980 to 2020, the local area's contribution evaporation accounted for 1.2% to 27.1% of the total contribution evaporation, while external sources provided between 60.4% and 93.3%. This distribution is significantly influenced by monsoonal patterns, which make external moisture the primary source to precipitation in the VMD. Seasonally, during the dry season, the amount of contribution evaporation within the precipitationshed ranges from 0.28 mm to 19.47 mm, whereas in the wet season, it increases to between 0.83 mm and 106.17 mm. The predominant moisture sources shift from the northeastern regions, like the South China Sea, in the dry seasons to the southwestern areas, such as the Bay of Bengal, in the wet seasons. The closer to the target region (i.e., the VMD), the more evaporation is contributed from these sources due to the effect of precipitation sinking during the moisture transport. A strong correlation exists between the contribution evaporation and ERA5 precipitation data, with the model tracking approximately 70% of the precipitation moisture from the precipitationsheds. Further, the negative anomalies in contribution evaporation closely align with historical drought events in the VMD. A marked moisture deficit from external sources constitutes the primary shortfall during drought years, underscoring the significance of external moisture in regional drought dynamics.

- Which factors predominantly influence moisture transport and how do they affect the occurrence and evolution of droughts in the VMD?

Correlation analysis is usually insufficient for deducing causal relationships among Earth system variables, necessitating more robust analytical methods. In this study, therefore, the PCMCI+ algorithm was utilized to establish the direction of causal links, while the CCM test assessed the strength of these relationships. This causality analysis is instrumental in identifying key drivers of atmospheric moisture transport that influence the onset and progression of droughts, which is essential for improving drought prediction models.

The results in Chapter 4 indicate that the causality network exhibits seasonal variability. During the dry season, TCW_external significantly affects the contribution evaporation, which in turn influences precipitation levels in the VMD. Conversely, in the wet season, wind speed emerges as the primary factor affecting the contribution evaporation from external sources. The constrained moisture flux during dry seasons, due to drier air conditions in the upwind areas, highlights humidity as a limiting factor. Therefore, TCW stands out as the critical variable restricting contribution evaporation in the dry season. In the wet season, however, with abundant water vapor in the upwind areas, anomalies in wind speed dominate the water vapor flux dynamics.

The causal link between contribution evaporation from external sources and precipitation exhibits a stronger relation in the dry season (0.87) compared to the wet season (0.60), underscoring its heightened relevance to the VMD precipitation during drier periods. Furthermore, the causal connections between TCW_local and precipitation suggest that local atmospheric conditions substantially impact precipitation volumes, with a stronger causal link observed during the dry season (0.71) than in the wet season (0.40).

In summary, the investigation of research questions in Chapter 4 highlighted the pivotal role of external moisture transport in the initiation and progression of droughts in the VMD. Additionally, this analysis elucidates the primary drivers of atmospheric moisture transport that govern drought dynamics. Such insights are crucial for enhancing the development of an accurate drought prediction model in the VMD.

Objective 2: To quantitatively assess the LA interactions in the VMD using deep learning techniques, as described in Chapter 5. Research questions focus on:

- Can deep learning algorithms effectively capture and simulate the LA interactions in the VMD?

Previous studies underscore the critical role of LA interactions in the development and intensification of droughts. Effective simulation of these interactions by deep learning algorithms (e.g., the LSTNet model) is essential for elucidating the mechanisms underlying these processes and highlighting the significance of key variables during drought periods in the VMD.

Comparative assessments reveal that the LSTNet model outperforms the RegCM model in simulating key climatic variables including precipitation, soil moisture, and sensible and latent heat. Specifically, LSTNet marginally underestimates precipitation with a BIAs of -0.082, compared to a more substantial underestimation by RegCM (BIAs = -0.286). Notably, RegCM also tends to overestimate extreme precipitation events, reaching values 1.6 times higher than those observed in ERA5 data. In contrast, LSTNet closely aligns with ERA5 measurements for soil moisture, sensible heat and latent heat, exhibiting strong correlations and minimal BIAs and outstripping the performance of RegCM. Both LSTNet and RegCM perform better in simulating continuous and periodic variables (i.e., soil moisture, sensible and latent heat) than the variables with a discrete distribution (e.g., precipitation).

The weight analysis reveals that LA interactions among these variables occurred within two days. Additionally, soil moisture, sensible and latent heat, wind speed and middle-level specific humidity play an important role in the simulation of precipitation. The persistence of soil moisture is shown to play a dominant role in the simulation of its own dynamics, with notable contributions from latent heat, sensible heat, and precipitation. Sensible heat is largely influenced by low-level U-wind and soil moisture, while latent heat is primarily governed by soil moisture, low-level V-wind, its persistence and precipitation. These interactions and the relative importance of each variable highlight the model's potential to uncover complex interactions among key variables, even in the absence of explicitly encoded physical principles.

Overall, the LSTNet neural network demonstrates robust performance in simulating LA interactions in the VMD. Therefore, LSTNet is well-suited for conducting quantitative analyses in LA interactions across the region.

- What is the extent of the influence exerted by key variables within the LA interactions?

The quantitative assessment of key variables in LA interactions is crucial for a deeper understanding of influential mechanisms and their role in exacerbating drought conditions in the VMD. As illustrated in Chapter 5, anomalies in sensible heat can decrease precipitation by up to 20% during dry periods and soil moisture exerts a significant influence on precipitation in both dry and wet conditions. The effects of sensible heat on precipitation may relate to its coupling with temperature and convective processes. In the context of global warming, sensible heat will increase and inhibit the formation of precipitation to some extent in the future. In addition, a declining trend in soil moisture is likely to reduce precipitation during dry periods.

Overall, the LSTNet neural network adeptly captures the relative importance of key variables like sensible heat in LA interactions across various climates, with outputs that align closely with physical processes. Evaluating the impact of sensible heat is vital as it can inhibit precipitation formation and intensify drought severity. In the context of climate change, the combined effects of decreased soil moisture and increased sensible heat are likely to further reduce precipitation, elevate temperatures, and enhance the probability of droughts. This study not only enriches understanding of influential mechanisms in LA interactions during drought evolution and intensification but also offers valuable insights for the further development and enhancement of hydrological models for drought monitoring and forecasting.

Objective 3: To develop a deep learning model that can accurately predict drought conditions in the VMD, as presented in Chapter 6. Research questions are listed below:

- How effectively can the deep learning algorithm predict drought conditions in the VMD?

An effective drought prediction model is essential for developing and utilizing drought early warning systems, which play a crucial role in implementing drought mitigation strategies and enhancing resilience and adaptability to droughts. In this context, a deep neural network was developed in this study, specifically the ConvGRU model, to provide accurate and reliable drought predictions in the VMD.

The ConvGRU model demonstrates strong performance in predicting meteorological and agricultural droughts across the four lead times. It successfully captures approximately 90%

of meteorological drought events and about 80% of agricultural drought events at a 3-month lead time. The model excels in both the POD and the FAR, with fewer than 10% of the drought months and events being falsely predicted at this lead time. Additionally, ConvGRU effectively identifies compound dry-hot conditions, capturing 67% of compound dry-hot months and 55% of severe compound dry-hot months at the 3-month lead time, and 83% of compound dry-hot events overall.

At the 3-month lead time, ConvGRU accurately predicted the onset of both meteorological and agricultural drought events with the ABias of approximately 0.7 months. The model maintains an ABias of less than two months for predicting drought duration, and ABias for intensity and severity are lower than 0.62 and 1.4, respectively. In terms of the predictions of representative drought events, ConvGRU effectively predicted the most severe meteorological drought and the longest agricultural drought at the lead time of 3 months but underestimated the severity of the onset stage of the most severe compound dry-hot event.

Overall, the ConvGRU neural network offers a robust tool for accurate and reliable drought prediction in the VMD, particularly with a lead time of 3 months, underpinning its potential utility in enhancing regional drought management strategies.

- Do atmospheric conditions from external precipitation source regions enhance the performance of the drought prediction model?

The ConvGRU model's enhanced performance in predicting droughts at the 3-month lead time can be attributed to the delayed influence of atmospheric conditions (i.e., specific humidity, U-wind, V-wind) from external areas through the process of moisture transport. Notably, the performance of neural networks typically declines as the lead time extends, but the ConvGRU model achieves optimal results at the 3-month interval.

Incorporating atmospheric data from the external precipitation source region substantially enhances the ConvGRU neural network's ability to forecast droughts in the VMD, especially for predictions at the lead time of 3 months. These findings highlight the crucial role of external atmospheric conditions in enhancing the accuracy of sub-seasonal to seasonal drought predictions. Such improvements are vital for the development of effective early warning systems, providing a crucial tool for anticipatory actions and management strategies.

In summary, as stated previously, droughts are generally linked with large-scale atmospheric circulation anomalies, such as moisture transport, and terrestrial processes like land-atmosphere interactions, with the former critical for drought initiation and the latter for their evolution and intensification. Understanding these moisture transport processes and LA interactions enhances comprehension of drought dynamics in the VMD, which in turn informs the development of more accurate and reliable prediction models such as the ConvGRU. The insights gained are fundamentally important for advancing early drought warning systems and mitigation strategies, as well as for strengthening drought resilience and adaptability in the VMD.

7.2 Limitations of the research

Despite the contributions highlighted above, this study still has some limitations.

7.2.1 Limitations in quantifying the effects of external humidity and wind on the VMD precipitation

In Chapter 4, the moisture tracking model WAM-2layers and causal inference algorithms, including PCMCI+ and CCM, were employed to investigate the influence of upwind atmospheric conditions on drought propagation in the VMD and to identify the principal factors driving moisture transport that impacts VMD precipitation. The causality analysis highlighted that specific humidity and wind speed in the upwind areas are the two primary drivers of contribution evaporation and precipitation in the VMD during dry and wet seasons, respectively. During the specific drought events, i.e., droughts in 2015-2016 and 2019-2020, the decreased moisture transport during the 2016 dry season was primarily due to anomalies in both humidity and wind speed. Conversely, the 2020 dry season's reduced moisture transport was predominantly influenced by humidity, while in the 2019 wet season, a wind speed anomaly was the main factor reducing moisture transport to the VMD.

However, this research didn't quantify the extent to which the contribution evaporation and VMD precipitation was affected by either specific humidity or wind speed in the upwind area. Even though the causal inference algorithms identified the dominant factors that drive the moisture transport processes and VMD precipitation, they are limited to quantifying the impacts of these factors. Because the input of causality algorithms is a spatial averaged multivariate time series, the causality algorithms capture the whole picture of how wind speed and humidity in the external area affect moisture transport.

As for the research of quantifying the effects of wind speed and humidity on moisture transport anomalies from the perspective of moisture transport processes, Benedict et al. (2021) and Guan et al. (2022) decomposed the integrated vertical moisture flux into two components: dynamic, which is dominated by wind speed, and thermodynamic, which is governed by specific humidity. Their findings indicate that during drought periods, moisture transport anomalies within the targeted study areas are predominantly driven by dynamic processes. Similarly, Yuan et al. (2021) observed that in two representative regions of China, the dynamic component largely governs the changes in moisture flux convergence through the convergence or divergence term. However, this method of decomposition has limitations. It quantifies the effects of wind speed and humidity on moisture transport anomalies specifically within the target study areas rather than in the upwind areas. This limitation arises because wind speed is a vector quantity, meaning its directional properties (e.g., eastward or westward for U-wind) significantly influence its interpretation. Further explanation in Chapter 4's discussion highlights that although a negative humidity anomaly was prevalent across much of the precipitationshed during the dry season of 2020, the thermodynamic component exhibited both positive and negative anomalies. This variability was due to contrasting climatological U-wind directions in the lower (westward) and upper (eastward) parts of the precipitationshed. Consequently, the anomalies in the thermodynamic component do not provide a clear quantitative measure of humidity's impact on moisture transport. Moreover, the impact of wind speed anomalies on moisture transportation can differ significantly between dry and wet seasons, subsequently affecting precipitation in the target area. For instance, a negative U-wind anomaly during the 2020 dry season facilitated increased moisture transport to the VMD, whereas a similar negative anomaly during the 2019 wet season obstructed moisture transport, exacerbating drought conditions in the VMD.

In this study, therefore, in addition to the analyses of causal effects of humidity and wind speed on the VMD precipitation, only the anomalies of TCW and U-wind were considered to analyse their effect on the recent two severe droughts. Quantifying the specific contributions of external humidity and wind speed to VMD precipitation remains challenging. This difficulty arises from the reliance on the atmospheric moisture budget equation used in the moisture tracking model (Van der Ent et al., 2014; Zhang, 2020; Guan et al., 2022). The amount of precipitation is hard to estimate under the conditions of climatological wind or humidity. Thus, it is difficult to track contribution evaporation directly with climatological wind or humidity and subsequently separate the effects of wind and specific humidity on the amount of tracked contribution evaporation quantitatively. Future research could profitably explore the distinct impacts of wind and specific humidity on the amounts of tracked contribution evaporation, enhancing understanding of these dynamic processes in climate modelling.

7.2.2 Limitations of the LSTNet neural network in simulating LA interactions globally

In Chapter 5, the potential applicability of the LSTNet neural network for simulating LA interactions beyond the VMD was discussed, specifically in Central California. It revealed that the LSTNet effectively captures the influence of key variables (e.g., sensible heat) on LA interactions within Central California. Similar to the findings in the VMD, it is projected that under global warming, sensible heat will increase, potentially reducing precipitation levels in this region in the future. Furthermore, soil moisture anomalies were found to negatively affect surface temperatures in both VMD and Central California, specifically, the drier the soil moisture, the higher the temperature. However, LSTNet was originally developed for the prediction of multivariate time series, therefore, all the variables within both the VMD and Central California were spatially averaged as multivariate time series. This requirement imposes a limitation, confining the applicability of the LSTNet to regional studies where hydrometeorological processes and properties are relatively uniform. The use of spatially averaged multivariate time series is inadequate for simulating LA interactions on larger scales, such as continental or global levels, due to the variability of LA interactions across different climate regions. For example, the influence of sensible heat anomalies on precipitation in Central California is somewhat less pronounced than in the VMD. In contrast, the impact of soil moisture on surface temperature is much stronger in Central California.

There are two methods to overcome this limitation to extend the applicability of the LSTNet neural network in simulating LA interactions at a continental or global scale. First, the research area (whether a continent or the globe) could be subdivided into smaller climate regions that share similar hydrometeorological characteristics. Each region could then have a specifically configured LSTNet model trained on its distinct spatially averaged data. Second, the LSTNet model could be integrated with other neural networks that are capable of processing spatial information, such as CNNs or GNNs. For example, Gao et al. (2022) and Bi et al. (2023) employed a 3D space-time transformer to forecast medium-range global weather, and Lam et al. (2023) developed a GNN-based neural network called GraphCast, which is a key advance in accurate and efficient global weather forecasting.

7.3 Perspectives on future research

Based on the analysis and investigations in this study, there are several future research directions and questions were identified, which are discussed below.

7.3.1 The validation of deep neural networks simulation with the GLACE-CMIP5 experiments

Previously, the Global Land-Atmosphere Coupling Experiment (GLACE) was initially established using 12 Atmospheric General Circulation Models (AGCMs) to quantify atmospheric responses to anomalies in land surface states, thereby assessing the strength of LA coupling (Koster et al., 2004, 2006; Guo et al., 2006). Moreover, given that soil moisture plays a key role in the LA interactions and climate systems, the GLACE-CMIP5 experiment was later conducted within the framework of CMIP5. This experiment aimed to explore the impact of soil moisture on long-term climatic changes under both historical and future scenarios (Seneviratne et al., 2013; Schwingshackl et al., 2018). In GLACE-CMIP5, control simulations maintained the original fully coupled soil moisture-climate interactions, while experimental simulations (expA and expB) utilized prescribed soil moisture based on the seasonal cycle of 1971-2000 climatology and a 30-year running mean, respectively. The difference between the control and experimental simulations allowed for precise quantification of soil moisture variability’s contributions to long-term changes in climate within LA interactions. The GLACE-CMIP5 framework has been extensively used to analyze soil moisture’s effects on temperature and precipitation extremes (Lorenz et al., 2016) and land carbon uptake (Green et al., 2019; Humphrey et al., 2021). Furthermore, studies indicate that soil moisture-atmosphere feedbacks may mitigate the decrease of water availability in arid regions (Zhou S. et al., 2021) and intensify future droughts and atmospheric aridity (Berg et al., 2016; Zhou S. et al., 2019). However, it specifically investigates how and to what extent the soil moisture variability affects the climate system (e.g., precipitation, temperature), but ignores the role of other important climate factors (e.g., sensible heat) in the system. In this study, the LSTNet neural network was employed to model LA interactions and to evaluate the influence of key variables, including but not limited to soil moisture, across the VMD and Central California regions. Even though the LSTNet can capture the relationships among those variables very well, it is still a deep learning algorithm rather than a physical process-based model. Therefore, it would be of great significance to validate the LSTNet outputs with the GLACE-CMIP5 model.

7.3.2 Is extensive long-term historical data beneficial to the development of deep learning models?

Earth is increasingly susceptible to a range of natural hazards, such as droughts, which are becoming more impactful due to heightened exposure in hazard-prone areas under the changing climate (Dai, 2011; Sheffield et al., 2012). This trend underscores a critical need to enhance predictive understanding of Earth's systems (Reidet al., 2010; Bergen et al., 2019; Reichstein et al., 2019; Sun et al., 2022). In the past few decades, geoscience has been experiencing a transformative data revolution. A large amount of data sourced from in-situ observations, remote sensing, and physics-based Earth system models are now accessible, offering unprecedented opportunities for scientific advancement (Faghmous et al., 2014; Srivastava et al., 2017; Camps-Valls et al., 2021). In the meantime, deep learning has emerged as a leading and rapidly evolving technique within Artificial Intelligence, playing a crucial role in extracting valuable information and offering new insights into the Earth system (Bergen et al., 2019; Munappy et al., 2019; Tahmasebi et al., 2020; Bailly et al., 2022). These techniques are pivotal in exploring hydrometeorological hazards, extreme climate events, and in advancing spatial-temporal predictions and Earth system modeling (Liu and Wu, 2016; Sharma et al., 2017; Racah et al., 2017; Gao et al., 2022; Bi et al., 2023; Zhang et al., 2023).

Despite the rapid expansion of in-situ and remote sensing hydrometeorological datasets, coupled with considerable advancements in deep learning algorithms, the effectiveness of employing extensive historical datasets in training deep learning models warrants careful consideration. Consequently, it is crucial to evaluate and discuss whether deep learning models systematically benefit from the inclusion of long-term historical data, because the deep learning models' performance depends critically on data quality (Whang et al., 2020, 2023). In general, a larger dataset can enhance a model's learning and generalization capabilities (Mårtensson et al., 2020; Sharma et al., 2020; Luca et al., 2022). However, if the data are of poor quality (e.g., non-representative, containing ambiguous values, noise, irrelevant features, and inconsistency), the efficacy of deep learning models may be significantly compromised, leading to diminished accuracy (Sarker, 2021). Consequently, this necessitates a detailed consideration of multiple factors influencing the application of historical data in machine learning frameworks.

First, data quality is critical when leveraging extensive datasets for deep learning applications. Although large datasets are typically considered beneficial, their effectiveness is contingent upon their quality (Fenza et al., 2021). Usually, these datasets are influenced by noise, irrelevant information, and errors, which, if not properly cleaned and processed, can deteriorate deep learning models' performance (Maranghi et al., 2022; Aldoseri et al., 2023). Taking the ERA5 dataset as an example, as indicated in Figure 6.11, there are significant reduction in the uncertainty in precipitation, soil moisture, and sensible and latent heat from

1940 to 1979. Similarly, this trend continues, as described by Hersbach et al. (2020), with the uncertainty in temperature, zonal wind, and specific humidity showing sustained declines from 1979 to 2018. This is mainly because the development and progress of measurement methods and processing algorithms have brought more reliable and accurate data (Hersbach et al., 2020). Therefore, although the historical depth of ERA5 data can potentially enrich deep learning model training, careful consideration is still necessary due to the elevated uncertainty in data recorded before 1979.

Second, timeliness and relevance of data are critical factors in the utility of datasets for training and building machine learning models. These attributes ensure that the data remain relevant to the current context and are appropriate and directly applicable to the specific tasks or problems being addressed (Batista and Monard, 2018; Aldoseri et al., 2023). Temporal dynamics in data, such as concept drift, can render historical data less relevant or even misleading the model over time (Tsymbal, 2004; Lu et al., 2018; Suárez-Cetrulo et al., 2023). For example, in sectors like finance or consumer behaviour, data patterns frequently shift in response to evolving market conditions or changing consumer preferences (Suárez-Cetrulo et al., 2019; Masegosa et al., 2020). Similarly, although hydrometeorological patterns do not alter abruptly, the patterns of drought processes and LA interactions have evolved over the past century and will continue to develop in the future with the changing climate (Loukas et al., 2008; Strzepek et al., 2010; Menéndez et al., 2016; Cook et al., 2018; Satoh et al., 2022). As indicated in Chapter 6, the ConvGRU model underestimates meteorological drought and compound dry-hot conditions after 2000 due to training on data spanning from 1940 to 1999, which fails to capture recent climatic change developments in the 21st century. Furthermore, in Chapter 5, future changes in the interactions among precipitation, soil moisture, and sensible and latent heat based on CMIP6 projections indicate evolving LA patterns. Therefore, while developing deep learning models with extensive long-term data, such as ERA5 data from 1940, it is important to assess whether these data accurately reflect contemporary climate characteristics.

Computational complexity is also a critical consideration when deploying large datasets to train deep learning models. The extensive computational resources required for handling very large datasets can significantly prolong training times, depending on the algorithmic complexity, may render some computations infeasible (Chen and Lin, 2014). The effect of data complexity on the expressive capabilities of deep learning models is still rarely explored (Hu et al., 2021). Taking the weather forecasting models GraphCast and

GraphCast_operational as an example, despite the GraphCast model processing data from 37 pressure levels, it is the GraphCast_operational model, with input from only 13 pressure levels, which demonstrates superior performance in predicting precipitation, even without precipitation data as input (Rasp et al., 2024). This suggests that an increase in the volume and complexity of historical data might not only extend training times but could also potentially impair the performance of deep neural networks.

In addition, the presence of extensive datasets can lead to overfitting in models, particularly when the data lack diversity or excessive emphasis is placed on detailing features that do not generalize well to unseen data (Dos Santos et al., 2009; Srivastava et al., 2014; Jabbar and Khan, 2015; Zhang et al., 2018).

Given the factors discussed above, the strategic utilization of large-scale historical data in the development of deep neural networks necessitates rigorous data management. Effective collection and preprocessing of data are essential to optimize the performance of these deep neural networks, ensuring that they not only fit the training data but also generalize well to new, unseen data (Najafabadi et al., 2015; Gheisari et al., 2017; Munappy et al., 2019, 2022; Aldoseri et al., 2023).

7.3.3 How to develop and implement the drought early warning system based on the drought predictions

The recurring nature of drought disasters around the world has underscored the necessity for robust disaster management strategies (Hao et al., 2017). Effective disaster management typically follows a continuous cycle comprising response, recovery, mitigation, and preparedness phases (Schmitt et al., 2007; UNISDR, 2009). These phases encompass actions taken before, during, and after a drought to mitigate its impacts and facilitate recovery. As illustrated in Masupha et al. (2021), the response phase involves immediate measures post-drought, such as distributing livestock feed in affected areas (Ng and Yap, 2011); recovery focuses on restoring physical and social systems, where successful outcomes often depend on subsequent seasonal conditions, such as above-normal precipitation to replenish soil water, while slow recovery may imply aggravated impacts (Khan et al., 2008; Ruehr et al., 2019); mitigation involves strategies to lessen drought impacts, including promoting effective water conservation (Van Zyl, 2006), while preparedness involves proactive measures to anticipate and manage potential drought impacts.

Drought early warning systems play a pivotal role in enhancing preparedness and mitigation efforts, providing critical alerts to government agencies, businesses, and communities about upcoming drought risks (Masupha et al., 2021; Sene, 2024). However, despite advances in drought prediction detailed in Chapter 6 through deep neural networks and external atmospheric data, significant challenges remain in developing and fully implementing drought early warning systems in the VMD. A drought early warning system extends far beyond simple forecasting, it serves as a comprehensive framework that supports effective response strategies through the collection, processing, analysis, and communication of critical drought-related information, along with actionable recommendations to mitigate potential damage (Pulwarty and Sivakumar, 2014; Masupha et al., 2021). According to the United Nations International Strategy for Disaster Reduction (UNISDR, 2006), such systems must be both people- and location-centered, encompassing four integral elements: (1) knowledge of the risks faced (Al Alhmoudi, 2016); (2) technical monitoring and warning service; (3) dissemination of meaningful warnings to those at risk; and (4) public awareness and preparedness to act (Basher, 2006). Similarly, the Sendai Framework (United Nations Office for Disaster Risk Reduction, 2015) emphasizes the need to (1) tailor early warning systems to meet user-specific needs, including social and cultural considerations; (2) promote the use of simple and cost-effective early warning equipment and infrastructure; and (3) expand dissemination channels for natural disaster early warning information. The Integrated Drought Management Programme (IDMP), a collaborative initiative between the World Meteorological Organization (WMO) and the Global Water Partnership (GWP), also highlights the necessity of timely communication of reliable drought condition information to water and land managers, policymakers, and the public through appropriate channels. Such communication serves as a critical foundation for reducing vulnerability and enhancing the mitigation and response capacities of both people and systems at risk (WMO and GWP, 2016).

Historically, drought early warning systems have evolved significantly since their inception in the 1980s, driven by the severe famines in Sudan and Ethiopia (Kim and Guha-Sapir, 2012). Initial systems were largely reactive, focusing on monitoring and response after the onset of droughts (Cowan et al., 2014). They primarily utilized drought indices based on meteorological data and field reports. However, the subjective and delayed availability of this data highlighted the necessity for more timely and objective forecasting methods (Kumar et al., 2021).

Over the years, drought early warning systems have undergone significant evolution, by integrating a broader spectrum of information and technologies, and enhancing their accessibility and utility (Kafle, 2017). Remote sensing technology, for instance, enables the continuous monitoring and detection of drought conditions over extensive areas (Bokusheva et al., 2016). Further enhancements in computer modeling and data analysis have broadened the scope of drought early warning systems to include variables such as climate variability, socio-economic conditions, and land use patterns (Masupha et al., 2024).

The 2000s witnessed marked improvements in the technical tools available for risk assessment, prediction, and warning dissemination, driven by a deepening understanding of natural hazards (Basher, 2006; UNISDR, 2006). The advent of the digital age introduced sophisticated information and communication technologies, enabling the development of web-based drought early warning systems. These systems leverage internet communication and mobile technologies, such as smartphones and laptops, enhancing their functionality with timely and effective dissemination of warnings and crucial information (Wu and Wilhite, 2004; Pozzi, 2013; Hao et al., 2014, 2017; Datta, 2023). Such advancements have significantly improved the implementation and effectiveness of drought early warning systems, making them pivotal in proactive disaster management and mitigation strategies.

The rapid development of web-based drought early warning systems reflects a global commitment to enhancing proactive drought management. For example, in the US, the National Integrated Drought Information System (NIDIS) and the National Drought Mitigation Center (NDMC) at the University of Nebraska collaborate to enhance drought impact assessments, refine forecasts, and develop watershed-scale information portals. (Western Governors' Association, 2004; Pulwarty and Sivakumar, 2014). This effort is supported by the partnerships across federal, state, regional, and local agencies, research institutions, and the private sector, ultimately establishing a national drought early warning system for diverse sectors (Owen et al., 2007; Umphlett et al., 2012; Masupha et al., 2021). The web portal (https://www.drought.gov/drought/) of this system provides essential data on current and forecasted drought conditions from sources like the NOAA Climate Prediction Center and the U.S. Drought Monitor (Svoboda et al., 2002; Masupha et al., 2021).

Another prominent initiative is the Famine Early Warning Systems Network (FEWS NET, http://www.fews.net/), which monitors food security across approximately 38 countries in sub-Saharan Africa, Central Asia, Central America, and the Caribbean, offering early warning information and collaborating with over 20 organizations to provide comprehensive updates on food security, disaster probabilities, climatic conditions, and market dynamics (Brown M.

E., 2008; Funk et al., 2019; Funk et al., 2019; Masupha et al., 2021). In Asia, the High Resolution South Asia Drought Monitor (https://sites.google.com/a/iitgn.ac.in/high_resolution_south_asia_drought_monitor/) developed by the Indian Institute of Technology - Gandhi Nagar (IIT-GN) and the International Water Management Institute (IWMI), offers real-time drought monitoring and forecasting for South Asia and on a national scale for countries including India, Pakistan, and Bangladesh, utilizing indices such as SPI, SSI, and SRI (Shah and Mishra, 2015; Aadhar and Mishra, 2017). Similarly, in New Zealand, the National Institute of Water and Atmospheric Research's New Zealand Drought Monitor system provides accessible drought information to drought-sensitive stakeholders like farmers, commercial consultants and government officials (Mol et al., 2017). There are also several widely used drought early warning systems around the world, such as the Middle East and North Africa Regional Drought Management Systems (MENA RDMS, Fragaszy et al., 2020) based on the composite drought indicator (CDI, Bijaber et al., 2018), and the Czech Drought Monitor System for the improvement of crop production decisions (Trnka et al., 2020). Collectively, these systems illustrate a sophisticated approach outlined by Otkin et al. (2022), which integrates a physically based identification framework, enhanced monitoring capabilities, and improved predictive accuracy across various time scales. This comprehensive strategy aids in impact assessments and guides policy, significantly bolstering the capacity of decision-makers to monitor, predict, plan for, and respond effectively to drought challenges.

As for the drought early warning systems in the VMD, Takeuchi et al. (2015) established a satellite-based, near-real-time drought monitoring system applicable to croplands across several Asian countries, including Cambodia, China, Myanmar, Laos, Thailand, Vietnam, and Indonesia. While this system delivers historical and near-real-time data on drought conditions, a truly effective early warning system relies on a comprehensive, multi-sectoral approach that encompasses collaboration across all stages of the warning process from monitoring and prediction to response and evaluation (Pulwarty and Sivakumar, 2014). On the other hand, a local drought early warning system could equip local farmers to better anticipate and prepare for drought conditions, despite potential limitations in their resources or capacity to adapt (Andersson et al., 2020). Therefore, given the VMD's heavy reliance on agriculture and aquaculture, developing and implementing a robust, effective drought early warning system based on accurate sub-seasonal to seasonal forecasts is crucial. Such a system would significantly enhance the region's ability to manage and mitigate the impacts of drought, safeguarding its agricultural productivity and broader economic stability.

7.4 Conclusion

This study explores the physical processes that govern drought dynamics in the VMD, from external atmospheric moisture transport to local LA interactions. Initially, the research findings from Chapter 4 demonstrate that the contribution evaporation from external sources significantly exceeds that from local sources, with seasonal monsoons influencing the primary moisture sources, shifting from northeastern areas during the dry season (northeast monsoon) to southwestern regions in the wet season (southwest monsoon). The causality analysis underscores the crucial roles of humidity and wind speed in the moisture transport process across different seasons, highlighting the importance of external atmospheric conditions in both the initiation and progression of droughts in the VMD. Subsequently, the influence of soil moisture, sensible heat, and latent heat on LA interactions was quantitatively assessed using the LSTNet neural network. LSTNet demonstrates strong performance in simulating LA interactions, effectively capturing complex relationships among key LA variables, even without explicitly incorporating physical principles. It reveals that soil moisture and sensible heat significantly affect precipitation levels in the VMD, particularly during dry periods. Moreover, a decline in soil moisture coupled with an increase in sensible heat is projected to elevate temperatures, reduce precipitation, and thereby enhance the likelihood of future droughts. Finally, in response to the pivotal role of external atmospheric conditions in the drought initiation, the ConvGRU deep neural network was developed and detailed in Chapter 6 to predict drought conditions in the VMD. The model effectively predicts sub-seasonal to seasonal drought conditions, including meteorological droughts, agricultural droughts and compound dry-hot events, particularly at the lead time of 3 months. The accuracy and reliability of the ConvGRU model form a solid basis for the development and application of drought early warning systems in the region. In summary, this research not only deepens understanding of the mechanisms behind drought dynamics but also pioneers an effective drought prediction model. Such advancements are crucial for enhancing drought resilience and adaptability in the VMD.

References

Aadhar, S., & Mishra, V. (2017). High-resolution near real-time drought monitoring in South Asia. *Scientific Data*, *4*(1), 1-14. https://doi.org/10.1038/sdata.2017.145

Achite, M., Gul, E., Elshaboury, N., Jehanzaib, M., Mohammadi, B., & Danandeh Mehr, A. (2023a). An improved adaptive neuro-fuzzy inference system for hydrological drought prediction in Algeria. *Physics and Chemistry of the Earth*, *131*(April), 103451. https://doi.org/10.1016/j.pce.2023.103451

Achite, M., Katipoğlu, O. M., Jehanzaib, M., Elshaboury, N., Kartal, V., & Ali, S. (2023b). Hydrological drought prediction based on hybrid extreme learning machine: Wadi Mina Basin case study, Algeria. *Atmosphere*, *14*(9), 1–23. https://doi.org/10.3390/atmos14091447

Achour, K., Meddi, M., Zeroual, A., Bouabdelli, S., Maccioni, P., & Moramarco, T. (2020). Spatio-temporal analysis and forecasting of drought in the plains of northwestern Algeria using the standardized precipitation index. *Journal of Earth System Science*, *129*(1). https://doi.org/10.1007/s12040-019-1306-3

Adamson, P. (2005). Drought study: analysis, forecasting, planning and management. *Mekong River Commission, Vientiane, Laos*.

AghaKouchak, A. (2014). A baseline probabilistic drought forecasting framework using standardized soil moisture index: Application to the 2012 United States drought. *Hydrology and Earth System Sciences*, *18*(7), 2485–2492. https://doi.org/10.5194/hess-18-2485-2014

AghaKouchak, A., Cheng, L., Mazdiyasni, O., & Farahmand, A. (2014). Global warming and changes in risk of concurrent climate extremes: Insights from the 2014 California drought. *Geophysical Research Letters*, *41*(24), 8847–8852. https://doi.org/10.1002/2014GL062308

AghaKouchak, A., Mirchi, A., Madani, K., Di Baldassarre, G., Nazemi, A., Alborzi, A., Anjileli, H., Azarderakhsh, M., Chiang, F., Hassanzadeh, E., Huning, L. S., Mallakpour, I., Martinez, A., Mazdiyasni, O., Moftakhari, H., Norouzi, H., Sadegh, M., Sadeqi, D., Van Loon, A. F., & Wanders, N. (2021). Anthropogenic drought: Definition, challenges, and opportunities. *Reviews of Geophysics*, *59*(2), 1–23. https://doi.org/10.1029/2019RG000683

AghaKouchak, A., Pan, B., Mazdiyasni, O., Sadegh, M., Jiwa, S., Zhang, W., Love, C. A., Madadgar, S., Papalexiou, S. M., Davis, S. J., Hsu, K., & Sorooshian, S. (2022). Status and prospects for drought forecasting: opportunities in artificial intelligence and hybrid physical-statistical forecasting. *Philosophical Transactions of the Royal Society A: Mathematical,* *Physical and Engineering Sciences*, *380*(2238). https://doi.org/10.1098/rsta.2021.0288

Aghelpour, P., Mohammadi, B., Biazar, S. M., Kisi, O., & Sourmirinezhad, Z. (2020). A theoretical approach for forecasting different types of drought simultaneously, using entropy theory and machine-learning methods. *ISPRS International Journal of Geo-Information*, *9*(12). https://doi.org/10.3390/ijgi9120701

Aghelpour, P., Mohammadi, B., Mehdizadeh, S., Bahrami-Pichaghchi, H., & Duan, Z. (2021). A novel hybrid dragonfly optimization algorithm for agricultural drought prediction. *Stochastic Environmental Research and Risk Assessment*, *35*(12), 2459–2477. https://doi.org/10.1007/s00477-021-02011-2

Al Alhmoudi, A. (2016). *Developing a framework for integrated community-centered early warning system to enhance disaster resilience in UAE*. University of Salford (United Kingdom). PhD Dissertations & Theses.

Al Mamun, M. A., Sarker, M. R., Sarkar, M. A. R., Roy, S. K., Nihad, S. A. I., McKenzie, A. M., Hossain, M. I., & Kabir, M. S. (2024). Identification of influential weather parameters and seasonal drought prediction in Bangladesh using machine learning algorithm. *Scientific Reports*, *14*(1), 1–21. https://doi.org/10.1038/s41598-023-51111-2

Alahacoon, N., & Edirisinghe, M. (2022). A comprehensive assessment of remote sensing and traditional based drought monitoring indices at global and regional scale. *Geomatics, Natural Hazards and Risk*, *13*(1), 762-799. https://doi.org/10.1080/19475705.2022.2044394

Aldoseri, A., Al-Khalifa, K. N., & Hamouda, A. M. (2023). Re-thinking data strategy and integration for artificial intelligence: concepts, opportunities, and challenges. *Applied Sciences*, *13*(12), 7082. https://doi.org/10.3390/app13127082

Alessi, M. J., Herrera, D. A., Evans, C. P., DeGaetano, A. T., & Ault, T. R. (2022). Soil moisture conditions determine land-atmosphere coupling and drought risk in the Northeastern United States. *Journal of Geophysical Research: Atmospheres*, *127*(6), 1–14. https://doi.org/10.1029/2021JD034740

Ali, S., Hasan, U., Li, X., Faruque, O., Sampath, A., Huang, Y., Gani, M. O., & Wang, J. (2024). *Causality for Earth Science -- A review on time-series and spatiotemporal causality methods*. 1–66. http://arxiv.org/abs/2404.05746

American Meteorological Society (AMS) (2004). Statement on meteorological drought. *Bulletin of the American Meteorological Society*, *85*, 771-773.

An, R., Zhang, L., Wang, Z., Quaye-Ballard, J. A., You, J., Shen, X., Gao, W., Huang, L. J., Zhao, Y., & Ke, Z. (2016). Validation of the ESA CCI soil moisture product in China. *International Journal of Applied Earth Observation and Geoinformation*, *48*, 28–36. https://doi.org/10.1016/j.jag.2015.09.009

Anderegg, W. R., Schwalm, C., Biondi, F., Camarero, J. J., Koch, G., Litvak, M., Ogle, K., Shaw, J. D., Shevliakova, E., Williams, A. P., Wolf, A., Ziaco, E., & Pacala, S. (2015). Pervasive drought legacies in forest ecosystems and their implications for carbon cycle models. *Science*, *349*(6247), 528–532. https://doi.org/10.1126/science.aab1833

Andersson, L., Wilk, J., Graham, L. P., Wikner, J., Mokwatlo, S., & Petja, B. (2020). Local early warning systems for drought–Could they add value to nationally disseminated seasonal climate forecasts? *Weather and Climate Extremes*, *28*, 100241. https://doi.org/10.1016/j.wace.2019.100241

Arneth, A., Mercado, L., Kattge, J., & Booth, B. B. B. (2012). Future challenges of representing land-processes in studies on land-atmosphere interactions. *Biogeosciences*, *9*(9), 3587–3599. https://doi.org/10.5194/bg-9-3587-2012

Asharaf, S., Dobler, A., & Ahrens, B. (2012). Soil moisture-precipitation feedback processes in the Indian summer monsoon season. *Journal of Hydrometeorology*, *13*(5), 1461–1474. https://doi.org/10.1175/JHM-D-12-06.1

Ault, T. R. (2020). On the essentials of drought in a changing climate. *Science*, *368*(6488), 256-260. https://doi.org/10.1126/science.aaz5492

Ayugi, B., Eresanya, E. O., Onyango, A. O., Ogou, F. K., Okoro, E. C., Okoye, C. O., Anoruo, C. M., Dike, V. N., Ashiru, O. R., Daramola, M. T., Mumo, R., & Ongoma, V. (2022). Review of meteorological drought in Africa: historical trends, impacts, mitigation measures, and prospects. *Pure and Applied Geophysics*, *179*(4), 1365-1386. https://doi.org/10.1007/s00024-022-02988-z

Bai, D., Lu, G., Zhu, Z., Zhu, X., Tao, C., Fang, J., & Li, Y. (2022). Prediction interval estimation of landslide displacement using bootstrap, variational mode decomposition, and long and short-term time-series network. *Remote Sensing*, *14*(22), 5808. https://doi.org/10.3390/rs14225808

Bai, J., Cui, Q., Chen, D., Yu, H., Mao, X., Meng, L., & Cai, Y. (2018). Assessment of the SMAP-derived soil water deficit index (SWDI-SMAP) as an agricultural drought index in

China. *Remote Sensing*, *10*(8). https://doi.org/10.3390/RS10081302

Bailly, A., Blanc, C., Francis, É., Guillotin, T., Jamal, F., Wakim, B., & Roy, P. (2022). Effects of dataset size and interactions on the prediction performance of logistic regression and deep learning models. *Computer Methods and Programs in Biomedicine*, *213*, 106504. https://doi.org/10.1016/j.cmpb.2021.106504

Baker, A. J., Sodemann, H., Baldini, J. U., Breitenbach, S. F., Johnson, K. R., van Hunen, J., & Zhang, P. (2015). Seasonality of westerly moisture transport in the East Asian summer monsoon and its implications for interpreting precipitation δ18O. *Journal of Geophysical Research: Atmospheres*, *120*(12), 5850-5862. https://doi.org/10.1002/2014JD022919

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw U, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., & Wofsy, S. (2001). FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, *82*(11), 2415–2434. https://doi.org/10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2

Banadkooki, F. B., Singh, V. P., & Ehteram, M. (2021). Multi-timescale drought prediction using new hybrid artificial neural network models. *Natural Hazards*, *106*(3), 2461–2478. https://doi.org/10.1007/s11069-021-04550-x

Barrett, A. B., Duivenvoorden, S., Salakpi, E. E., Muthoka, J. M., Mwangi, J., Oliver, S., & Rowhani, P. (2020). Forecasting vegetation condition for drought early warning systems in pastoral communities in Kenya. *Remote Sensing of Environment*, *248*(June 2019). https://doi.org/10.1016/j.rse.2020.111886

Başakın, E. E., Ekmekcioğlu, Ö., & Özger, M. (2021). Drought prediction using hybrid soft-computing methods for semi-arid region. *Modeling Earth Systems and Environment*, *7*, 2363-2371. https://doi.org/10.1007/s40808-020-01010-6

Basher, R. (2006). Global early warning systems for natural hazards: systematic and people-centred. *Philosophical transactions of the royal society a: mathematical, physical and engineering sciences*, *364*(1845), 2167-2182. https://doi.org/10.1098/rsta.2006.1819

Bastakoti, R. C., Gupta, J., Babel, M. S., & van Dijk, M. P. (2014). Climate risks and adaptation strategies in the Lower Mekong River basin. *Regional Environmental Change*, *14*(1), 207-219. https://doi.org/10.1007/s10113-013-0485-8

Batista, G. E., & Monard, M. C. (2018). Data quality in machine learning: A study in the context of imbalanced data. *Neurocomputing*, *275*, 1665-1679.

Beamesderfer, E. R., Buechner, C., Faiola, C., Helbig, M., Sanchez-Mejia, Z. M., Yáñez-Serrano, A. M., Zhang, Y., & Richardson, A. D. (2022). Advancing cross-disciplinary understanding of land-atmosphere interactions. *Journal of Geophysical Research: Biogeosciences*, *127*(2), 1–16. https://doi.org/10.1029/2021JG006707

Belayneh, A., & Adamowski, J. (2012). Standard precipitation index drought forecasting using neural networks, wavelet neural networks, and support vector regression. *Applied Computational Intelligence and Soft Computing*, *2012*, 1–13. https://doi.org/10.1155/2012/794061

Belayneh, A., Adamowski, J., Khalil, B., & Ozga-Zielinski, B. (2014). Long-term SPI drought forecasting in the Awash River Basin in Ethiopia using wavelet neural networks and wavelet support vector regression models. *Journal of Hydrology*, *508*, 418–429. https://doi.org/10.1016/j.jhydrol.2013.10.052

Belayneh, A., Adamowski, J., Khalil, B., & Quilty, J. (2016). Coupling machine learning methods with wavelet transforms and the bootstrap and boosting ensemble approaches for drought prediction. *Atmospheric Research*, *172–173*, 37–47. https://doi.org/10.1016/j.atmosres.2015.12.017

Benedict, I., van Heerwaarden, C. C., van der Linden, E. C., Weerts, A. H., & Hazeleger, W. (2021). Anomalous moisture sources of the Rhine basin during the extremely dry summers of 2003 and 2018. *Weather and Climate Extremes*, *31*(December 2020), 100302. https://doi.org/10.1016/j.wace.2020.100302

Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*, *5*(2), 157-166. https://doi.org/10.1109/72.279181

Berg, A., Findell, K., Lintner, B., Giannini, A., Seneviratne, S. I., Van Den Hurk, B., Lorenz, R., Pitman, A., Hagemann, S., Meier, A., Cheruy, F., Ducharne, A., Malyshev, S., & Milly, P. C. D. (2016). Land-atmosphere feedbacks amplify aridity increase over land under global warming. *Nature Climate Change*, *6*(9), 869–874. https://doi.org/10.1038/nclimate3029

Bergen, K. J., Johnson, P. A., de Hoop, M. V., & Beroza, G. C. (2019). Machine learning for data-driven discovery in solid Earth geoscience. *Science*, *363*(6433), eaau0323. https://doi.org/10.1126/science.aau0323

Beringer, J., Hutley, L. B., McHugh, I., Arndt, S. K., Campbell, D., Cleugh, H. A., Cleverly, J., Resco de Dios, V., Eamus, D., Evans, B., Ewenz, C., Grace, P., Griebel, A., Haverd, V., Hinko-Najera, N., Huete, A., Isaac, P., Kanniah, K., Leuning, R., Liddell, M. J., Macfarlane, C., Meyer, W., Moore, C., Pendall, E., Phillips, A., Phillips, R. L., Prober, S. M., Restrepo-Coupe, N., Rutledge, S., Schroder, I., Silberstein, R., Southall, P., Yee, M. S., Tapper, N. J., van Gorsel, E., Vote, C., Walker, J., & Wardlaw, T (2016). An introduction to the Australian and New Zealand flux tower network - OzFlux. *Biogeosciences*, *13*(21), 5895–5916. https://doi.org/10.5194/bg-13-5895-2016

Berman, J. D., Ebisu, K., Peng, R. D., Dominici, F., & Bell, M. L. (2017). Drought and the risk of hospital admissions and mortality in older adults in western USA from 2000 to 2013: a retrospective study. *The Lancet Planetary Health*, *1*(1), e17–e25. https://doi.org/10.1016/S2542-5196(17)30002-5

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., & Tian, Q. (2023). Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, *619*(7970), 533–538. https://doi.org/10.1038/s41586-023-06185-3

Bijaber, N., El Hadani, D., Saidi, M., Svoboda, M. D., Wardlow, B. D., Hain, C. R., Poulsen, C. C., Yessef, M., & Rochdi, A. (2018). Developing a remotely sensed drought monitoring indicator for Morocco. *Geosciences*, *8*(2), 55. https://doi.org/10.3390/geosciences8020055

Bisselink, B., & Dolman, A. J. (2008). Precipitation recycling: Moisture sources over Europe using ERA-40 data. *Journal of Hydrometeorology*, *9*(5), 1073–1083. https://doi.org/10.1175/2008JHM962.1

Boken, V. K., Cracknell, A. P., & Heathcote, R. L. (2005). *Monitoring and predicting agricultural drought: a global study*. Oxford University Press.

Bokusheva, R., Kogan, F., Vitkovskaya, I., Conradt, S., & Batyrbayeva, M. (2016). Satellite-based vegetation health indices as a criteria for insuring against drought-related yield losses. *Agricultural and Forest Meteorology*, *220*, 200-206. https://doi.org/10.1016/j.agrformet.2015.12.066

Bolinger, R. A., Gronewold, A. D., Kompoltowicz, K., & Fry, L. M. (2017). Application of the nmme in the development of a new regional seasonal climate forecast tool. *Bulletin of the*

American Meteorological Society, *98*(3), 555–564. https://doi.org/10.1175/BAMS-D-15-00107.1

Bonaccorso, B., Cancelliere, A., & Rossi, G. (2015). Probabilistic forecasting of drought class transitions in Sicily (Italy) using Standardized Precipitation Index and North Atlantic Oscillation Index. *Journal of Hydrology*, *526*, 136–150. https://doi.org/10.1016/j.jhydrol.2015.01.070

Bonavita, M., Hólm, E., Isaksen, L., & Fisher, M. (2016). The evolution of the ECMWF hybrid data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *142*(694), 287–303. https://doi.org/10.1002/qj.2652

Bosilovich, M. G., & Schubert, S. D. (2002). Water vapor tracers as diagnostics of the regional hydrologic cycle. *Journal of Hydrometeorology*, *3*(2), 149-165. https://doi.org/10.1175/1525-7541(2002)003<0149:WVTADO>2.0.CO;2

Bosilovich, M. G., Chen, J., Robertson, F. R., & Adler, R. F. (2008). Evaluation of global precipitation in reanalyses. *Journal of Applied Meteorology and Climatology*, *47*(9), 2279–2299. https://doi.org/10.1175/2008JAMC1921.1

Breshears, D. D., Cobb, N. S., Rich, P. M., Price, K. P., Allen, C. D., Balice, R. G., Romme, W. H., Kastens, J. H., Floyd, M. L., Belnap, J., Anderson, J. J., Myers, O. B., & Meyer, C. W. (2005). Regional vegetation die-off in response to global-change-type drought. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(42), 15144–15148. https://doi.org/10.1073/pnas.0505734102

Brito, C. S. de, Silva, R. M. da, Santos, C. A. G., Brasil Neto, R. M., & Coelho, V. H. R. (2021). Monitoring meteorological drought in a semiarid region using two long-term satellite-estimated rainfall datasets: A case study of the Piranhas River basin, northeastern Brazil. *Atmospheric Research*, *250*(July 2020). https://doi.org/10.1016/j.atmosres.2020.105380

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., Dorigo, W., Matgen, P., Martínez-Fernández, J., Llorens, P., Latron, J., Martin, C., & Bittelli, M. (2011). Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe. *Remote Sensing of Environment*, *115*(12), 3390–3408. https://doi.org/10.1016/j.rse.2011.08.003

Brocca, L., Melone, F., Moramarco, T., Wagner, W., & Hasenauer, S. (2010). ASCAT soil wetness index validation through in situ and modeled soil moisture data in central Italy.

Remote Sensing of Environment, *114*(11), 2745–2755. https://doi.org/10.1016/j.rse.2010.06.009

Brown, J. F., Wardlow, B. D., Tadesse, T., Hayes, M. J., & Reed, B. C. (2008). The Vegetation Drought Response Index (VegDRI): A new integrated approach for monitoring drought stress in vegetation. *GIScience and Remote Sensing*, *45*(1), 16–46. https://doi.org/10.2747/1548-1603.45.1.16

Brown, M. E., 2008: *Famine Early Warning Systems and Remote Sensing Data*. Springer, 314 pp., https://doi.org/10.1007/978-3-540-75369-8.

Brubaker, K. L., Entekhabi, D., & Eagleson, P. S. (1993). Estimation of continental precipitation recycling. *Journal of Climate*, *6*(6), 1077-1089. https://doi.org/10.1175/1520-0442(1993)006%3C1077:EOCPR%3E2.0.CO;2

Burde, G. I., & Zangvil, A. (2001). The estimation of regional precipitation recycling. Part I: Review of recycling models. *Journal of climate*, *14*(12), 2497-2508. https://doi.org/10.1175/1520-0442(2001)014%3C2497:TEORPR%3E2.0.CO;2

Butterworth, B. J., Desai, A. R., Townsend, P. A., Petty, G. W., Andresen, C. G., Bertram, T. H., Kruger, E. L., Mineau, J. K., Olson, E. R., Paleri, S., Pertzborn, R. A., Pettersen, C., Stoy, P. C., Thom, J. E., Vermeuel, M. P., Wagner, T. J., Wright, D. B., Zheng, T., Metzger, S., Schwartz, M. D., Iglinski, T. J., Mauder, M., Speidel, J., Vogelmann, H., Wanner, L., Augustine, T. J., Brown, W. O. J., Oncley, S. P., Buban, M., Lee, T. R., Cleary, P., Durden, D. J., Florian, C. R., Lantz, K., Riihimaki, L. D., Sedlar, J., Meyers, T. P., Plummer, D. M., Guzman, E. R., Smith, E. N., Sühring, M., Turner, D. D., Wang, Z., White, L. D., & Wilczak, J. M. (2021). Connecting land–atmosphere interactions to surface heterogeneity in CHEESEHEAD19. *Bulletin of the American Meteorological Society*, *102*(2), E421–E445. https://doi.org/10.1175/BAMS-D-19-0346.1

Cammalleri, C., Micale, F., & Vogt, J. (2016). A novel soil moisture-based drought severity index (DSI) combining water deficit magnitude and frequency. *Hydrological Processes*, *30*(2), 289–301. https://doi.org/10.1002/hyp.10578

Camps-Valls, G., Tuia, D., Zhu, X. X., & Reichstein, M. (Eds.). (2021). *Deep learning for the Earth Sciences: A comprehensive approach to remote sensing, climate science and geosciences*. John Wiley & Sons.

Cao, J., Guan, F., Zhang, X., Nam, W. H., Leng, G., Gao, H., Ye, Q., Gu, X., Zeng, J., Zhang, X., Huang, T., & Niyogi, D. (2023). Multiple Markov chains for categorial drought prediction on the US drought monitor at weekly scale. *Journal of Applied Meteorology and Climatology*, *62*(10), 1415-1435. https://doi.org/10.1175/JAMC-D-23-0061.1

Chahine, M. T. (1992). GEWEX: The global energy and water cycle experiment. *Eos, Transactions American Geophysical Union*, *73*(2), 9-14. https://doi.org/10.1029/91EO00007

Chen, J., Huang, W., Feng, S., Zhang, Q., Kuang, X., Chen, J., & Chen, F. (2021). The modulation of westerlies-monsoon interaction on climate over the monsoon boundary zone in East Asia. *International Journal of Climatology*, *41*, E3049-E3064. https://doi.org/10.1002/joc.6903

Chen, X. W., & Lin, X. (2014). Big data deep learning: challenges and perspectives. *IEEE access*, *2*, 514-525. https://doi.org/10.1109/ACCESS.2014.2325029

Cheng, K., Wei, Z., Li, X., & Ma, L. (2024). Multi-source dataset assessment and variation characteristics of snow depth in Eurasia from 1980 to 2018. *Atmosphere*, *15*(5), 530. https://doi.org/10.3390/atmos15050530

Cheng, L., & AghaKouchak, A. (2015). A methodology for deriving ensemble response from multimodel simulations. *Journal of Hydrology*, *522*, 49–57. https://doi.org/10.1016/j.jhydrol.2014.12.025

Cheng, T. F., & Lu, M. (2020). Moisture source-receptor network of the East Asian summer monsoon land regions and the associated atmospheric steerings. *Journal of Climate*, *33*(21), 9213–9231. https://doi.org/10.1175/JCLI-D-19-0868.1

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*. https://doi.org/10.48550/arXiv.1406.1078

Choi, M., Jacobs, J. M., Anderson, M. C., & Bosch, D. D. (2013). Evaluation of drought indices via remotely sensed data with hydrological variables. *Journal of Hydrology*, *476*, 265–273. https://doi.org/10.1016/j.jhydrol.2012.10.042

Christian, J. I., Basara, J. B., Hunt, E. D., Otkin, J. A., Furtado, J. C., Mishra, V., Xiao, X., & Randall, R. M. (2021). Global distribution, trends, and drivers of flash drought occurrence. *Nature Communications*, *12*(1), 1–11. https://doi.org/10.1038/s41467-021-26692-z

Christian, J. I., Basara, J. B., Otkin, J. A., Hunt, E. D., Wakefield, R. A., Flanagan, P. X., &

Xiao, X. (2019). A methodology for flash drought identification: Application of flash drought frequency across the United States. *Journal of Hydrometeorology*, *20*(5), 833–846. https://doi.org/10.1175/JHM-D-18-0198.1

Chu, Q. C., Wang, Q. G., & Feng, G. L. (2017). Determination of the major moisture sources of cumulative effect of torrential rain events during the preflood season over South China using a Lagrangian particle model. *Journal of Geophysical Research: Atmospheres*, *122*(16), 8369–8382. https://doi.org/10.1002/2016JD026426

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*. https://doi.org/10.48550/arXiv.1412.3555

Ciric, D., Stojanovic, M., Drumond, A., Nieto, R., & Gimeno, L. (2016). Tracking the origin of moisture over the danube river basin using a Lagrangian approach. *Atmosphere*, *7*(12), 1–12. https://doi.org/10.3390/atmos7120162

Colliander, A., Jackson, T. J., Bindlish, R., Chan, S., Das, N., Kim, S. B., Cosh, M. H., Dunbar, R. S., Dang, L., Pashaian, L., Asanuma, J., Aida, K., Berg, A., Rowlandson, T., Bosch, D., Caldwell, T., Caylor, K., Goodrich, D., al Jassar, H., Lopez-Baeza, E., Martínez-Fernández, J., González-Zamora, A., Livingston, S., McNairn, H., Pacheco, A., Moghaddam, M., Montzka, C., Notarnicola, C., Niedrist, G., Pellarin, T., Prueger, J., Pulliainen, J., Rautiainen, K., Ramos, J., Seyfried, M., Starks, P., Su, Z., Zeng, Y., van der Velde, R., Thibeault, M., Dorigo, W., Vreugdenhil, M., Walker, J. P., Wu, X., Monerris, A., O'Neill, P. E., Entekhabi, D., Njoku, E.G., & Yueh, S. (2017). Validation of SMAP surface soil moisture products with core validation sites. *Remote Sensing of Environment*, *191*, 215–231. https://doi.org/10.1016/j.rse.2017.01.021

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., BroNnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., Worley, S. J. (2011). The Twentieth Century Reanalysis Project. *Quarterly Journal of the Royal Meteorological Society*, *137*(654), 1–28. https://doi.org/10.1002/qj.776

Cook, B. I., Mankin, J. S., & Anchukaitis, K. J. (2018). Climate change and drought: From past to future. *Current Climate Change Reports*, *4*(2), 164–179. https://doi.org/10.1007/s40641-018-0093-2

Coppola, E., Stocchi, P., Pichelli, E., Torres Alavez, J. A., Glazer, R., Giuliani, G., Di Sante, F., Nogherotto, R., & Giorgi, F. (2021). Non-Hydrostatic RegCM4 (RegCM4-NH): Model description and case studies over multiple domains. *Geoscientific Model Development*, *14*(12), 7705–7723. https://doi.org/10.5194/gmd-14-7705-2021

Cosslett, T. L., & Cosslett, P. D. (2018). The Lower Mekong Basin: Rice production, climate change, ENSO, and Mekong dams. In *sustainable development of rice and water resources in mainland Southeast Asia and Mekong River Basin* (pp. 85-114). Springer Singapore. https://doi.org/10.1007/978-981-10-5613-0

Cowan, Y., O'Brien, E., & Rakotomalala-Rakotondrandria, N. (2014). Community-based early warning systems: key practices for DRR implementers. http://www.fao.org/documents/card/en/c/57c0a761-8d31-4f0f-8a07-27e1ea0568d5

Crausbay, S. D., Ramirez, A. R., Carter, S. L., Cross, M. S., Hall, K. R., Bathke, D. J., Betancourt, J. L., Colt, S., Cravens, A. E., Dalton, M. S., Dunham, J. B., Hay, L. E., Hayes, M. J., McEvoy, J., McNutt, C. A., Moritz, M. A., Nislow, K. H., Raheem, N., & Sanford, T. (2017). Defining ecological drought for the twenty-first century. *Bulletin of the American Meteorological Society*, *98*(12), 2543–2550. https://doi.org/10.1175/BAMS-D-16-0292.1

Dai, A. (2011). Drought under global warming: A review. *Wiley Interdisciplinary Reviews: Climate Change*, *2*(1), 45–65. https://doi.org/10.1002/wcc.81

Dai, A. (2013). Increasing drought under global warming in observations and models. *Nature Climate Change*, *3*(1), 52–58. https://doi.org/10.1038/nclimate1633

Dai, H., Kozareva, Z., Dai, B., Smola, A., & Song, L. (2018). Learning steady-states of iterative algorithms over graphs. In *International conference on machine learning* (pp. 1106-1114). PMLR.

Dai, M., Huang, S., Huang, Q., Leng, G., Guo, Y., Wang, L., Fang, W., Li, P., & Zheng, X. (2020). Assessing agricultural drought risk and its dynamic evolution characteristics. *Agricultural Water Management*, *231*(December 2019), 106003. https://doi.org/10.1016/j.agwat.2020.106003

Danabasoglu & Gokhan (2019). NCAR CESM2 model output prepared for CMIP6 ScenarioMIP ssp585 [Dataset]. *Earth System Grid Federation*. http://doi.org/10.22033/ESGF/CMIP6.7768

Danandeh Mehr, A., Rikhtehgar Ghiasi, A., Yaseen, Z. M., Sorman, A. U., & Abualigah, L.

(2023). A novel intelligent deep learning predictive model for meteorological drought forecasting. *Journal of Ambient Intelligence and Humanized Computing*, *14*(8), 10441–10455. https://doi.org/10.1007/s12652-022-03701-7

Dang, A., Vu, T. H., & Wang, J. C. (2017). A survey of deep learning for polyphonic sound event detection. In *2017 International Conference on Orange Technologies (ICOT)* (pp. 75-78). IEEE. https://doi.org/10.1109/ICOT.2017.8336092

Datta, P. (2023). The promise and challenges of the fourth industrial revolution (4IR). *Journal of Information Technology Teaching Cases*, *13*(1), 2-15. https://doi.org/10.1177/20438869211056938

Davy, R., & Esau, I. (2016). Differences in the efficacy of climate forcings explained by variations in atmospheric boundary layer depth. *Nature Communications*, *7*(1), 1-8. https://doi.org/10.1038/ncomms11690

Dayan, U., Nissen, K., & Ulbrich, U. (2015). Review Article: Atmospheric conditions inducing extreme precipitation over the eastern and western Mediterranean. *Natural Hazards and Earth System Sciences*, *15*(11), 2525–2544. https://doi.org/10.5194/nhess-15-2525-2015

De Arellano, J. V. G., Van Heerwaarden, C. C., & Lelieveld, J. (2012). Modelled suppression of boundary-layer clouds by plants in a CO 2 -rich atmosphere. *Nature Geoscience*, *5*(10), 701–704. https://doi.org/10.1038/ngeo1554

Delforge, D., Wathelet, V., Below, R., Sofia, C. L., Tonnelier, M., van Loenhout, J., & Speybroeck, N. (2023). EM-DAT: The emergency events database, Preprint. https://doi.org/10.21203/rs.3.rs-3807553/v1

Deyle, E. R., & Sugihara, G. (2011). Generalized theorems for nonlinear state space reconstruction. *PLoS ONE*, *6*(3). https://doi.org/10.1371/journal.pone.0018295

Dick, S. (2019). Artificial Intelligence. Harvard Data Science Review, 1(1). https://doi.org/10.1162/99608f92.92fe150c

Diem, P. K., Diem, N. K., Nguyen, C. T., & Minh, V. Q. (2024). Impacts of extreme drought on rice planting calendar in Vietnamese Mekong Delta. *Paddy and Water Environment*, *22*(1), 139-153.

Diez-Sierra, J., & del Jesus, M. (2020). Long-term rainfall prediction using atmospheric synoptic patterns in semi-arid climates with statistical and machine learning methods. *Journal of Hydrology*, *586*(March). https://doi.org/10.1016/j.jhydrol.2020.124789

Diffenbaugh, N. S., Singh, D., Mankin, J. S., Horton, D. E., Swain, D. L., Touma, D., Charland, A., Liu, Y., Haugen, M., Tsiang, M., & Rajaratnam, B. (2017). Quantifying the influence of global warming on unprecedented extreme climate events. *Proceedings of the National Academy of Sciences*, *114*(19), 4881-4886. https://doi.org/10.1073/pnas.1618082114

Diffenbaugh, N. S., Swain, D. L., Touma, D., & Lubchenco, J. (2015). Anthropogenic warming has increased drought risk in California. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(13), 3931–3936. https://doi.org/10.1073/pnas.1422385112

Dikshit, A., & Pradhan, B. (2021). Interpretable and explainable AI (XAI) model for spatial drought prediction. *Science of the Total Environment*, *801*, 149797. https://doi.org/10.1016/j.scitotenv.2021.149797

Dikshit, A., Pradhan, B., & Alamri, A. M. (2020a). Short-term spatio-temporal drought forecasting using random forests model at New South Wales, Australia. *Applied Sciences (Switzerland)*, *10*(12). https://doi.org/10.3390/app10124254

Dikshit, A., Pradhan, B., & Alamri, A. M. (2020b). Temporal hydrological drought index forecasting for New South Wales, Australia using machine learning approaches. *Atmosphere, 11*(6), 1–13. https://doi.org/10.3390/atmos11060585

Dikshit, A., Pradhan, B., & Alamri, A. M. (2021b). Long lead time drought forecasting using lagged climate variables and a stacked long short-term memory model. *Science of the Total Environment*, *755*, 142638. https://doi.org/10.1016/j.scitotenv.2020.142638

Dikshit, A., Pradhan, B., & Huete, A. (2021a). An improved SPEI drought forecasting approach using the long short-term memory neural network. *Journal of Environmental Management*, *283*(January), 111979. https://doi.org/10.1016/j.jenvman.2021.111979

Dikshit, A., Pradhan, B., & Santosh, M. (2022). Artificial neural networks in drought prediction in the 21st century–A scientometric analysis. *Applied Soft Computing*, *114*, 108080. https://doi.org/10.1016/j.asoc.2021.108080

Ding, Y., Hayes, M. J., & Widhalm, M. (2011). Measuring economic impacts of drought: A review and discussion. *Disaster Prevention and Management*, *20*(4), 434–446. https://doi.org/10.1108/09653561111161752

Ding, Y., Sun, Y., Wang, Z., Zhu, Y., & Song, Y. (2009). Inter-decadal variation of the summer precipitation in China and its association with decreasing Asian summer monsoon

Part II: Possible causes. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *29*(13), 1926-1944. https://doi.org/10.1002/joc.1759

Dirmeyer, P. A., Balsamo, G., Blyth, E. M., Morrison, R., & Cooper, H. M. (2021). Land-atmosphere interactions exacerbated the drought and heatwave over northern Europe during summer 2018. *AGU Advances*, *2*(2), 1–16. https://doi.org/10.1029/2020av000283

Dirmeyer, P. A., Jin, Y., Singh, B., & Yan, X. (2013). Trends in land-atmosphere interactions from CMIP5 simulations. *Journal of Hydrometeorology*, *14*(3), 829–849. https://doi.org/10.1175/JHM-D-12-0107.1

Dirmeyer, P. A., Wei, J., Bosilovich, M. G., & Mocko, D. M. (2014). Comparing evaporative sources of terrestrial precipitation and their extremes in MERRA using relative entropy. *Journal of Hydrometeorology*, *15*(1), 102–116. https://doi.org/10.1175/JHM-D-13-053.1

Djerbouai, S., & Souag-Gamane, D. (2016). Drought forecasting using neural networks, wavelet neural networks, and stochastic models: Case of the Algerois Basin in North Algeria. *Water Resources Management*, *30*(7), 2445–2464. https://doi.org/10.1007/s11269-016-1298-6

Dominguez, F., Kumar, P., Liang, X. Z., & Ting, M. (2006). Impact of atmospheric moisture storage on precipitation recycling. *Journal of Climate*, *19*(8), 1513–1530. https://doi.org/10.1175/JCLI3691.1

Dominguez, F., Miguez-Macho, G., & Hu, H. (2016). WRF with water vapor tracers: A study of moisture sources for the North American Monsoon. *Journal of Hydrometeorology*, *17*(7), 1915–1927. https://doi.org/10.1175/JHM-D-15-0221.1

Dong, W., Lin, Y., Wright, J. S., Xie, Y., Yin, X., & Guo, J. (2019). Precipitable water and CAPE dependence of rainfall intensities in China. *Climate Dynamics*, *52*, 3357-3368.

Dong, Z., Liu, H., Hu, H., Khan, M. Y. A., Wen, J., Chen, L., & Tian, F. (2022). Future projection of seasonal drought characteristics using CMIP6 in the Lancang-Mekong River Basin. *Journal of Hydrology*, *610*(April), 127815. https://doi.org/10.1016/j.jhydrol.2022.127815

Dorigo, W. A., Gruber, A., De Jeu, R. A. M., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R. M., & Kidd, R. (2015). Evaluation of the ESA CCI soil moisture product using ground-based observations. *Remote Sensing of Environment*, *162*, 380–395. https://doi.org/10.1016/j.rse.2014.07.023

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., Van Oevelen, P., Robock, A., & Jackson, T. (2011b). The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences*, *15*(5), 1675–1698. https://doi.org/10.5194/hess-15-1675-2011

Dorigo, W. A., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiova, A., Sanchis-Dufau, A. D., Zamojski, D., Cordes, C., Wagner, W., & Drusch, M. (2013). Global automated quality control of in situ soil moisture data from the International Soil Moisture Network. *Vadose Zone Journal*, *12*(3), vzj2012-0097. https://doi.org/10.2136/vzj2012.0097

Dorigo, W., Van Oevelen, P., Wagner, W., Drusch, M., Mecklenburg, S., Robock, A., & Jackson, T. (2011a). A new international network for in situ soil moisture data. *Eos*, *92*(17), 141–142. https://doi.org/10.1029/2011EO170001

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T., Lecomte, P. (2017). ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment*, *203*, 185–215. https://doi.org/10.1016/j.rse.2017.07.001

Dos Santos, E. M., Sabourin, R., & Maupin, P. (2009). Overfitting cautious selection of classifier ensembles with genetic algorithms. Information Fusion, 10(2), 150-162. https://doi.org/10.1016/j.inffus.2008.11.003

Doughty, C. E., Metcalfe, D. B., Girardin, C. A. J., Amézquita, F. F., Cabrera, D. G., Huasco, W. H., Silva-Espejo, J. E., Araujo-Murakami, A., Da Costa, M. C., Rocha, W., Feldpausch, T. R., Mendoza, A. L. M., Da Costa, A. C. L., Meir, P., Phillips, O. L., & Malhi, Y. (2015). Drought impact on forest carbon dynamics and fluxes in Amazonia. *Nature*, *519*(7541), 78–82. https://doi.org/10.1038/nature14213

Drumond, A., Gimeno, L., Nieto, R., Trigo, R. M., & Vicente-Serrano, S. M. (2017). Drought episodes in the climatological sinks of the Mediterranean moisture source: The role of moisture transport. *Global and Planetary Change*, *151*, 4–14. https://doi.org/10.1016/j.gloplacha.2016.12.004

Drumond, A., Muñiz, R. N., & Presa, L. G. (2016). A Lagrangian approach for investigating anomalies in the moisture transport during drought episodes. *Cuadernos de investigación geográfica: Geographical Research Letters*, *42*(1), 113-125. https://doi.org/10.18172/cig.2925

Drumond, A., Nieto, R., Hernandez, E., & Gimeno, L. (2011). A Lagrangian analysis of the variation in moisture sources related to drier and wetter conditions in regions around the Mediterranean Basin. *Natural Hazards and Earth System Sciences*, *11*(8), 2307–2320. https://doi.org/10.5194/nhess-11-2307-2011

Drumond, A., Stojanovic, M., Nieto, R., Vicente-Serrano, S. M., & Gimeno, L. (2019). Linking anomalous moisture transport and drought episodes in the IPCC reference regions. *Bulletin of the American Meteorological Society*, *100*(8), 1481–1498. https://doi.org/10.1175/BAMS-D-18-0111.1

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote sensing of Environment*, *120*, 25-36. https://doi.org/10.1016/j.rse.2011.11.026

Durre, I., Menne, M. J., & Vose, R. S. (2008). Strategies for evaluating quality assurance procedures. *Journal of Applied Meteorology and Climatology*, *47*(6), 1785–1791. https://doi.org/10.1175/2007JAMC1706.1

Durre, I., Menne, M. J., Gleason, B. E., Houston, T. G., & Vose, R. S. (2010). Comprehensive automated quality assurance of daily surface observations. *Journal of Applied Meteorology and Climatology*, *49*(8), 1615–1633. https://doi.org/10.1175/2010JAMC2375.1

Dutra, E., Pozzi, W., Wetterhall, F., Di Giuseppe, F., Magnusson, L., Naumann, G., Barbosa, P., Vogt, J., & Pappenberger, F. (2014). Global meteorological drought-Part 2: Seasonal forecasts. *Hydrology and Earth System Sciences*, *18*(7), 2669–2678. https://doi.org/10.5194/hess-18-2669-2014

Ebi, K. L., & Bowen, K. (2016). Extreme events as sources of health vulnerability: Drought as an example. *Weather and Climate Extremes*, *11*, 95–102. https://doi.org/10.1016/j.wace.2015.10.001

Edwards, D. C., & McKee, T. B. (1997). *Characteristics of 20 th century drought in the United States at multiple time scales* (Vol. 97, p. 155). Fort Collins: Colorado State University.

Ek, M. B., & Holtslag, A. A. M. (2004). Influence of soil moisture on boundary layer cloud development. *Journal of Hydrometeorology*, *5*(1), 86–99. https://doi.org/10.1175/1525-7541(2004)005<0086:IOSMOB>2.0.CO;2

El Ibrahimi, A., & Baali, A. (2018). Application of several artificial intelligence models for forecasting meteorological drought using the standardized precipitation index in the Saïss Plain (Northern Morocco). *International Journal of Intelligent Engineering & Systems*, *11*(1). https://doi.org/10.22266/ijies2018.0228.28

Emanuel, K. A. (1991). A scheme for representing cumulus convection in large-scale models. *Journal of Atmospheric Sciences, 48*(21), 2313-2329. https://doi.org/10.1175/1520-0469(1991)048<2313:ASFRCC>2.0.CO;2

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T. J., Johnson, J., Kimball, J., Piepmeier, J. R., Koster, R. D., Martin, N., McDonald, K. C., Moghaddam, M., Moran, S., Reichle, R., Shi, J. C., Spencer, M. W., Thurman, S. W., Tsang, L., & Van Zyl, J. (2010). The soil moisture active passive (SMAP) mission. *Proceedings of the IEEE*, *98*(5), 704–716. https://doi.org/10.1109/JPROC.2010.2043918

Esfahanian, E., Nejadhashemi, A. P., Abouali, M., Adhikari, U., Zhang, Z., Daneshvar, F., & Herman, M. R. (2017). Development and evaluation of a comprehensive drought index. *Journal of Environmental Management*, *185*, 31–43. https://doi.org/10.1016/j.jenvman.2016.10.050

Eslamian, S., Ostad-Ali-Askari, K., Singh, V. P., Dalezios, N. R., Ghane, M., Yihdego, Y., & Matouq, M. (2017). A review of drought indices. *Int. J. Constr. Res. Civ. Eng*, *3*(4). http://dx.doi.org/10.20431/2454-8693.0304005

European Commission (2007). Communication from the Commission to the European Parliament and the Council on addressing the challenge of water scarcity and droughts in the European Union. COM (2007) 414 final, Brussels

Faghmous, J. H., & Kumar, V. (2014). A big data guide to understanding climate change: The case for theory-guided data science. *Big data*, *2*(3), 155-163. https://doi.org/10.1089/big.2014.0026

Feng, S., Wu, X., Hao, Z., Hao, Y., Zhang, X., & Hao, F. (2020). A database for characteristics and variations of global compound dry and hot events. *Weather and Climate Extremes, 30*, 100299. https://doi.org/10.1016/j.wace.2020.100299

Fenza, G., Gallo, M., Loia, V., Orciuoli, F., & Herrera-Viedma, E. (2021). Data set quality in machine learning: consistency measure based on group decision making. *Applied Soft Computing*, *106*, 107366. https://doi.org/10.1016/j.asoc.2021.107366

Findell, K. L., & Eltahir, E. A. B. (2003). Atmospheric controls on soil moisture-boundary layer interactions. Part I: Framework development. *Journal of Hydrometeorology*, *4*(3), 552–569. https://doi.org/10.1175/1525-7541(2003)004<0552:ACOSML>2.0.CO;2

Fischer, T., Gemmer, M., Lüliu, L., & Buda, S. (2011). Temperature and precipitation trends and dryness/wetness pattern in the Zhujiang River Basin, South China, 1961-2007. *Quaternary International*, *244*(2), 138–148. https://doi.org/10.1016/j.quaint.2010.08.010

Ford, T. W., Rapp, A. D., Quiring, S. M., & Blake, J. (2015). Soil moisture–precipitation coupling: Observations from the Oklahoma Mesonet and underlying physical mechanisms. *Hydrology and Earth System Sciences*, 19(8), 3617-3631. https://doi.org/10.5194/hess-19-3617-2015

Fragaszy, S. R., Jedd, T., Wall, N., Knutson, C., Fraj, M. B., Bergaoui, K., Svoboda, M., Hayes, M., & McDonnell, R. (2020). Drought monitoring in the Middle East and North Africa (MENA) region: Participatory engagement to inform early warning systems. *Bulletin of the American Meteorological Society*, *101*(7), E1148-E1173. https://doi.org/10.1175/BAMS-D-18-0084.1

Frappart, F., Biancamaria, S., Normandin, C., Blarel, F., Bourrel, L., Aumont, M., Azemar, P., Vu, P. L., Le Toan, T., Lubac, B., & Darrozes, J. (2018). Influence of recent climatic events on the surface water storage of the Tonle Sap Lake. *Science of the Total Environment*, *636*, 1520–1533. https://doi.org/10.1016/j.scitotenv.2018.04.326

Fujiwara, M., Wright, J. S., Manney, G. L., Gray, L. J., Anstey, J., Birner, T., Davis, S., Gerber, E. P., Harvey, V. L., Hegglin, M. I., Homeyer, C. R., Knox, J. A., Krüger, K., Lambert, A., Long, C. S., Martineau, P., Molod, A., Monge-Sanz, B. M., Santee, M. L., Tegtmeier, S., Chabrillat, S., Tan, D. G. H., Jackson, D. R., Polavarapu, S., Compo, G. P., Dragani, R., Ebisuzaki, W., Harada, Y., Kobayashi, C., McCarty, W., Onogi, K., Pawson, S., Simmons, A., Wargan, K., Whitaker, J. S., & Zou, C.-Z. (2017) Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems. *Atmospheric Chemistry and Physics*, *17*(2), 1417–1452. https://doi.org/10.5194/acp-17-1417-2017

Fung, K. F., Huang, Y. F., Koo, C. H., & Soh, Y. W. (2020). Drought forecasting: A review of modelling approaches 2007–2017. *Journal of Water and Climate Change*, *11*(3), 771–799. https://doi.org/10.2166/wcc.2019.236

Funk, C., Shukla, S., Thiaw, W. M., Rowland, J., Hoell, A., McNally, A., Husak, G., Novella, N., Budde, M., Peters-Lidard, C., Adoum, A., Galu, G., Korecha, D., Magadzire, T., Rodriguez, M., Robjhon, M., Bekele, E., Arsenault, K., Peterson, P., Harrison, L., Fuhrman, S., Davenport, F., Landsfeld, M., Pedreros, D., Jacob, J. P., Reynolds, C., Becker-Reshef, I., & Verdin, J. (2019). Recognizing the Famine Early Warning Systems Network: Over 30 years of drought early warning science advances and partnerships promoting global food security. *Bulletin of the American Meteorological Society*, *100*(6), 1011-1027. https://doi.org/10.1175/BAMS-D-17-0233.1

Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning* (pp. 1050-1059). PMLR.

Gallicchio, C., & Micheli, A. (2010). Graph echo state networks. In *The 2010 international joint conference on neural networks (IJCNN)* (pp. 1-8). IEEE. https://doi.org/10.1109/IJCNN.2010.5596796

Gangl, M. (2010). Causal inference in sociological research. *Annual review of sociology*, *36*(1), 21-47. https://doi.org/10.1146/annurev.soc.012809.102702

Ganguli, P., & Janga Reddy, M. (2014). Ensemble prediction of regional droughts using climate inputs and the SVM-copula approach. *Hydrological Processes*, *28*(19), 4989–5009. https://doi.org/10.1002/hyp.9966

Gao, X., Shi, Y., Han, Z., Wang, M., Wu, J., Zhang, D., Xu, Y., & Giorgi, F. (2017). Performance of RegCM4 over major river basins in China. *Advances in Atmospheric Sciences*, *34*(4), 441–455. https://doi.org/10.1007/s00376-016-6179-7

Gao, Y., Chen, F., Miguez-Macho, G., & Li, X. (2020). Understanding precipitation recycling over the Tibetan Plateau using tracer analysis with WRF. *Climate Dynamics*, *55*(9–10), 2921–2937. https://doi.org/10.1007/s00382-020-05426-9

Gao, Z., Shi, X., Wang, H., Zhu, Y., Wang, Y., Li, M., & Yeung, D. Y. (2022). Earthformer: Exploring space-time transformers for Earth System forecasting. *Advances in Neural Information Processing Systems*, *35*(NeurIPS).

García-Herrera, R., Garrido-Perez, J. M., Barriopedro, D., Ordóñez, C., Vicente-Serrano, S.

M., Nieto, R., Gimeno, L., Sorí, R., & Yiou, P. (2019). The European 2016/17 drought. *Journal of Climate*, *32*(11), 3169–3187. https://doi.org/10.1175/JCLI-D-18-0331.1

Gawlikowski, J., Tassi, C. R. N., Ali, M., Lee, J., Humt, M., Feng, J., Kruspe, A., Triebel, R., Jung, P., Roscher, R., Shahzad, M., Yang, W., Bamler, R., & Zhu, X. X. (2023). A survey of uncertainty in deep neural networks. In *Artificial Intelligence Review* (Vol. 56, Issue s1). Springer Netherlands. https://doi.org/10.1007/s10462-023-10562-9

Geirinhas, J. L., Russo, A., Libonati, R., Sousa, P. M., Miralles, D. G., & Trigo, R. M. (2021). Recent increasing frequency of compound summer drought and heatwaves in Southeast Brazil. *Environmental Research Letters*, *16*(3), 034036. https://doi.org/10.1088/1748-9326/abe0eb

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., & Zhao, B. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419-5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Gentine, P., Holtslag, A. A. M., D'Andrea, F., & Ek, M. (2013). Surface and atmospheric controls on the onset of moist convection over land. *Journal of Hydrometeorology*, *14*(5), 1443–1462. https://doi.org/10.1175/JHM-D-12-0137.1

Gheisari, M., Wang, G., & Bhuiyan, M. Z. A. (2017). A survey on deep learning in big data. In *2017 IEEE international conference on computational science and engineering (CSE) and IEEE international conference on embedded and ubiquitous computing (EUC)* (Vol. 2, pp. 173-180). IEEE. https://doi.org/10.1109/CSE-EUC.2017.215

Gibon, F., Mialon, A., Richaume, P., Rodríguez-Fernández, N., Aberer, D., Boresch, A., Crapolicchio, R., Dorigo, W., Gruber, A., Himmelbauer, I., Preimesberger, W., Sabia, R., Stradiotti, P., Tercjak, M., & Kerr, Y. H. (2024). Estimating the uncertainties of satellite derived soil moisture at global scale. *Science of Remote Sensing*, *10*(July). https://doi.org/10.1016/j.srs.2024.100147

Gimeno, L., Drumond, A., Nieto, R., Trigo, R. M., & Stohl, A. (2010). On the origin of continental precipitation. *Geophysical Research Letters*, *37*(13), 1–7.

https://doi.org/10.1029/2010GL043712

Gimeno, L., Nieto, R., Drumond, A., Castillo, R., & Trigo, R. (2013). Influence of the intensification of the major oceanic moisture sources on continental precipitation. *Geophysical Research Letters*, *40*(7), 1443–1450. https://doi.org/10.1002/grl.50338

Gimeno, L., Stohl, A., Trigo, R. M., Dominguez, F., Yoshimura, K., Yu, L., Drumond, A., Durn-Quesada, A. M., & Nieto, R. (2012). Oceanic and terrestrial sources of continental precipitation. *Reviews of Geophysics*, *50*(4), 1–41. https://doi.org/10.1029/2012RG000389

Gimeno, L., Vázquez, M., Eiras-Barca, J., Sorí, R., Stojanovic, M., Algarra, I., Nieto, R., Ramos, A. M., Durán-Quesada, A. M., & Dominguez, F. (2020). Recent progress on the sources of continental precipitation as revealed by moisture transport analysis. *Earth-Science Reviews*, *201*(September 2019), 103070. https://doi.org/10.1016/j.earscirev.2019.103070

Gimeno-Sotelo, L., Sorí, R., Nieto, R., Vicente-Serrano, S. M., & Gimeno, L. (2024). Unravelling the origin of the atmospheric moisture deficit that leads to droughts. *Nature Water*, *2*(3), 242–253. https://doi.org/10.1038/s44221-023-00192-4

Giorgi, F., Coppola, E., Solmon, F., Mariotti, L., Sylla, M. B., Bi, X., Elguindi, N., Diro, G. T., Nair, V., Giuliani, G., Turuncoglu, U. U., Cozzini, S., Güttler, I., O'Brien, T. A., Tawfik, A. B., Shalaby, A., Zakey, A. S., Steiner, A. L., Stordal, F., Sloan, L. C., & Brankovic, C. (2012). RegCM4: Model description and preliminary tests over multiple CORDEX domains. *Climate Research*, *52*(1), 7–29. https://doi.org/10.3354/cr01018

Giorgi, F., Marinucci, M. R., Bates, G. T., & De Canio, G. (1993). Development of a Second-Generation Regional Climate Model (RegCM2). Part II: Convective processes and assimilation of lateral boundary conditions. *Monthly Weather Review*, *121*(10), 2814-2832. https://doi.org/10.1175/1520-0493(1993)121<2814:DOASGR>2.0.CO;2

Glass, T. A., Goodman, S. N., Hernán, M. A., & Samet, J. M. (2013). Causal inference in public health. *Annual review of public health*, *34*(1), 61-75. https://doi.org/10.1146/annurev-publhealth-031811-124606

Gómez-Hernández, M., Drumond, A., Gimeno, L., & Garcia-Herrera, R. (2013). Variability of moisture sources in the Mediterranean region during the period 1980-2000. *Water Resources Research*, *49*(10), 6781–6794. https://doi.org/10.1002/wrcr.20538

Gori, M., Monfardini, G., & Scarselli, F. (2005). A new model for learning in graph domains. In *Proceedings. 2005 IEEE international joint conference on neural networks, 2005.* (Vol. 2, pp. 729-734). IEEE. https://doi.org/10.1109/IJCNN.2005.1555942

Grabar, V., Marusov, A., Maximov, Y., Sotiriadi, N., Bulkin, A., & Zaytsev, A. (2024). Long-term drought prediction using deep neural networks based on geospatial weather data. *Environmental Modelling & Software*, 106127.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*(3), 424–438. https://doi.org/10.2307/1912791

Green, J. K., Konings, A. G., Alemohammad, S. H., Berry, J., Entekhabi, D., Kolassa, J., Lee, J. E., & Gentine, P. (2017). Regionally strong feedbacks between the atmosphere and terrestrial biosphere. *Nature Geoscience*, *10*(6), 410–414. https://doi.org/10.1038/ngeo2957

Green, J. K., Seneviratne, S. I., Berg, A. M., Findell, K. L., Hagemann, S., Lawrence, D. M., & Gentine, P. (2019). Large influence of soil moisture on long-term terrestrial carbon uptake. *Nature*, *565*(7740), 476–479. https://doi.org/10.1038/s41586-018-0848-x

Gruber, A., Scanlon, T., Van Der Schalie, R., Wagner, W., & Dorigo, W. (2019). Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology. *Earth System Science Data*, *11*(2), 717–739. https://doi.org/10.5194/essd-11-717-2019

Gu, Y., Hunt, E., Wardlow, B., Basara, J. B., Brown, J. F., & Verdin, J. P. (2008). Evaluation of MODIS NDVI and NDWI for vegetation drought monitoring using Oklahoma Mesonet soil moisture data. *Geophysical Research Letters*, *35*(22), 1–5. https://doi.org/10.1029/2008GL035772

Guan, Y., Gu, X., Slater, L. J., Li, L., Kong, D., Liu, J., Zhang, X., & Yan, X. (2022). Tracing anomalies in moisture recycling and transport to two record-breaking droughts over the Mid-to-Lower Reaches of the Yangtze River. *Journal of Hydrology*, *609*(March), 127787. https://doi.org/10.1016/j.jhydrol.2022.127787

Gudmundsson, L., & Seneviratne, S. I. (2015). European drought trends. *Proceedings of the International Association of Hydrological Sciences*, *369*, 75–79. https://doi.org/10.5194/piahs-369-75-2015

Gudmundsson, L., & Seneviratne, S. I. (2016). Anthropogenic climate change affects meteorological drought risk in Europe. *Environmental Research Letters*, *11*(4). https://doi.org/10.1088/1748-9326/11/4/044005

Guillod, B. P., Orlowsky, B., Miralles, D., Teuling, A. J., Blanken, P. D., Buchmann, N., Ciais, P., Ek, M., Findell, K. L., Gentine, P., Lintner, B. R., Scott, R. L., Van Den Hurk, B., &

Seneviratne, S. I. (2014). Land-surface controls on afternoon precipitation diagnosed from observational data: Uncertainties and confounding factors. *Atmospheric Chemistry and Physics*, *14*(16), 8343–8367. https://doi.org/10.5194/acp-14-8343-2014

Guo, H., Bao, A., Liu, T., Ndayisaba, F., He, D., Kurban, A., & De Maeyer, P. (2017). Meteorological drought analysis in the Lower Mekong Basin using satellite-based long-term CHIRPS product. *Sustainability (Switzerland)*, *9*(6). https://doi.org/10.3390/su9060901

Guo, H., John, J. G., Blanton, C., McHugh, C., Nikonov, S., Radhakrishnan, A., Rand, K., Zadeh, N. T., Balaji, V., Durachta, J., Dupuis, C., Menzel, R., Robinson, T., Underwood, S., Vahlenkamp, H., Dunne, K., Gauthier, P. PG, Ginoux, P., Griffies, S. M., Hallberg, R., Harrison, M., Hurlin, W., Lin, P., Malyshev, S., Naik, V., Paulot, F., Paynter, D, J., Ploshay, J., Schwarzkopf, D. M., Seman, C. J., Shao, A., Silvers, L., Wyman, B., Yan, X., Zeng, Y., Adcroft, A., Dunne, J. P., Held, I. M.,Krasting, J, P., Horowitz, L. W., Milly, C., Shevliakova, E., Winton, M., Zhao, M., Zhang, R. (2018). NOAA-GFDL GFDL-CM4 model output prepared for CMIP6 ScenarioMIP ssp585. [Dataset]. *Earth System Grid Federation*. https://doi.org/10.22033/ESGF/CMIP6.9268

Guo, R., Cheng, L., Li, J., Hahn, P. R., & Liu, H. (2020). A survey of learning causality with data: problems and methods. *ACM Computing Surveys*, *53*(4). https://doi.org/10.1145/3397269

Guo, Y., Huang, S., Huang, Q., Wang, H., Fang, W., Yang, Y., & Wang, L. (2019). Assessing socioeconomic drought based on an improved Multivariate Standardized Reliability and Resilience Index. *Journal of Hydrology*, *568*(November 2018), 904–918. https://doi.org/10.1016/j.jhydrol.2018.11.055

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S. (2016). Deep learning for visual understanding: A review. *Neurocomputing*, *187*, 27–48. https://doi.org/10.1016/j.neucom.2015.09.116

Guo, Z., Dirmeyer, P. A., Koster, R. D., Sud, Y. C. , Bonan, G., Oleson, K. W., Chan, E., Verseghy, D., Cox, P., Gordon, C. T. , McGregor, J. L. , Kanae, S., Kowalczyk, E., Lawrence, D., Liu, P., Mocko, D., Lu, C., Mitchell, K., Malyshev, S., McAvaney, B., Oki, T., Yamada, T., Pitman, A., Taylor, C. M., Vasic, R., & Xue, Y. (2006). GLACE: The Global Land–Atmosphere Coupling Experiment. Part II: Analysis. *Journal of Hydrometeorology*, *7*(4), 611-625. https://doi.org/10.1175/JHM511.1

Ha, T. V., Huth, J., Bachofer, F., & Kuenzer, C. (2022). A review of Earth observation-based drought studies in Southeast Asia. *Remote Sensing*, *14*(15), 1–30. https://doi.org/10.3390/rs14153763

Ha, T. V., Uereyen, S., & Kuenzer, C. (2023). Agricultural drought conditions over mainland Southeast Asia: Spatiotemporal characteristics revealed from MODIS-based vegetation time-series. *International Journal of Applied Earth Observation and Geoinformation*, *121*, 103378. https://doi.org/10.1016/j.jag.2023.103378

Haile, G. G., Tang, Q., Li, W., Liu, X., & Zhang, X. (2020). Drought: Progress in broadening its understanding. *Wiley Interdisciplinary Reviews: Water*, *7*(2), 1–25. https://doi.org/10.1002/WAT2.1407

Han, Z., Huang, S., Huang, Q., Leng, G., Wang, H., Bai, Q., Zhao, J., Ma, L., Wang, L., & Du, M. (2019). Propagation dynamics from meteorological to groundwater drought and their possible influence factors. *Journal of Hydrology*, *578*(August), 124102. https://doi.org/10.1016/j.jhydrol.2019.124102

Hao, Z., & AghaKouchak, A. (2013). Multivariate Standardized Drought Index: A parametric multi-index model. *Advances in Water Resources*, *57*, 12–18. https://doi.org/10.1016/j.advwatres.2013.03.009

Hao, Z., AghaKouchak, A., Nakhjiri, N., & Farahmand, A. (2014). Global integrated drought monitoring and prediction system. *Scientific data*, *1*(1), 1-10. https://doi.org/10.1038/sdata.2014.1

Hao, Z., Hao, F., Singh, V. P., & Zhang, X. (2018b). Changes in the severity of compound drought and hot extremes over global land areas. *Environmental Research Letters*, *13*(12). https://doi.org/10.1088/1748-9326/aaee96

Hao, Z., Singh, V. P., & Xia, Y. (2018a). Seasonal drought prediction: Advances, challenges, and future prospects. *Reviews of Geophysics*, *56*(1), 108–141. https://doi.org/10.1002/2016RG000549

Hao, Z., Yuan, X., Xia, Y., Hao, F., & Singh, V. P. (2017). An overview of drought monitoring and prediction systems at regional and global scales. *Bulletin of the American Meteorological Society*, *98*(9), 1879-1896. https://doi.org/10.1175/BAMS-D-15-00149.1

Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *International Journal of*

Climatology, *34*(3), 623–642. https://doi.org/10.1002/joc.3711

Hartmann, D. L., Klein Tank, A. M. G., Rusticucci, M., Alexander, L. V., Brönnimann, S., Charabi, Y. A. R., Dentener, F. J., Dlugokencky, E. J., Easterling, D. R., Kaplan, A., Soden, B. J., Thorne, P. W., Wild, M., & Zhai, P. (2013). Observations: Atmosphere and surface. In *Climate Change 2013 the Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Vol. 9781107057999). https://doi.org/10.1017/CBO9781107415324.008

Hawkins, E., & Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, *90*(8), 1095-1108.

Hayes, M. J., Svoboda, M. D., Wiihite, D. A., & Vanyarkho, O. V. (1999). Monitoring the 1996 drought using the standardized precipitation index. *Bulletin of the American meteorological society*, *80*(3), 429-438. https://doi.org/10.1175/1520-0477(1999)080%3C0429:MTDUTS%3E2.0.CO;2

Hayes, M., Svoboda, M., Wall, N., & Widhalm, M. (2011). The Lincoln declaration on drought indices: universal meteorological drought index recommended. *Bulletin of the American Meteorological Society*, *92*(4), 485-488. http://www.jstor.org/stable/26226865

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

Hecht, J. S., Lacombe, G., Arias, M. E., Dang, T. D., & Piman, T. (2019). Hydropower dams of the Mekong River basin: A review of their hydrological impacts. *Journal of Hydrology*, *568*(October 2018), 285–300. https://doi.org/10.1016/j.jhydrol.2018.10.045

Heim Jr, R. R. (2002). A review of twentieth-century drought indices used in the United States. *Bulletin of the American Meteorological Society*, *83*(8), 1149-1166. https://doi.org/10.1175/1520-0477-83.8.1149

Herrera-Estrada, J. E., Martinez, J. A., Dominguez, F., Findell, K. L., Wood, E. F., & Sheffield, J. (2019). Reduced moisture transport linked to drought propagation across North America. *Geophysical Research Letters*, *46*(10), 5243–5253. https://doi.org/10.1029/2019GL082475

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., & Thépaut, J-

N. (2023a). ERA5 hourly data on single levels from 1940 to present [Dataset]. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. https://doi.org/10.24381/cds.adbb2d47

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., & Thépaut, J-N. (2023b). ERA5 hourly data on pressure levels from 1940 to present [Dataset]. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. https://doi.org/10.24381/cds.bd0915c6

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G. D., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., & Thépaut, J. N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hilker, T., Lyapustin, A. I., Tucker, C. J., Sellers, P. J., Hall, F. G., & Wang, Y. (2012). Remote sensing of tropical ecosystems: Atmospheric correction and cloud masking matter. *Remote Sensing of Environment*, *127*, 370-384. https://doi.org/10.1016/j.rse.2012.08.035

Hirschi, M., Seneviratne, S. I., Alexandrov, V., Boberg, F., Boroneant, C., Christensen, O. B., Formayer, H., Orlowsky, B., & Stepanek, P. (2011). Observational evidence for soil-moisture impact on hot extremes in southeastern Europe. *Nature Geoscience*, *4*(1), 17–21. https://doi.org/10.1038/ngeo1032

Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *6*(02), 107-116. https://doi.org/10.1142/S0218488598000094

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Holgate, C. M., Van Dijk, A. I. J. M., Evans, J. P., & Pitman, A. J. (2020). Local and remote drivers of southeast Australian drought. *Geophysical Research Letters*, *47*(18), 1-10. https://doi.org/10.1029/2020GL090238

Holland, P. W. (1988). Causal inference, path analysis and recursive structural equations models. *ETS Research Report Series*, *1988*(1), i-50. https://doi.org/10.1002/j.2330-8516.1988.tb00270.x

Hollmann, R., Merchant, C. J., Saunders, R., Downy, C., Buchwitz, M., Cazenave, A., Chuvieco, E., Defourny, P., De Leeuw, G., Forsberg, R., Holzer-Popp, T., Paul, F., Sandven, S., Sathyendranath, S., Van Roozendael, M., & Wagner, W. (2013). The ESA climate change initiative: Satellite data records for essential climate variables. *Bulletin of the American Meteorological Society*, *94*(10), 1541–1552. https://doi.org/10.1175/BAMS-D-11-00254.1

Holtslag, A. A. M., De Bruijn, E. I. F., & Pan, H. L. (1990). A high resolution air mass transformation model for short-range weather forecasting. *Monthly Weather Review*, *118*(8), 1561-1575.

Horridge, M., Madden, J., & Wittwer, G. (2005). The impact of the 2002-2003 drought on Australia. *Journal of Policy Modeling*, *27*(3), 285–308. https://doi.org/10.1016/j.jpolmod.2005.01.008

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., & Iguchi, T. (2014). The global precipitation measurement mission. *Bulletin of the American Meteorological Society*, *95*(5), 701–722. https://doi.org/10.1175/BAMS-D-13-00164.1

Hou, A. Y., Skofronick-Jackson, G., Kummerow, C. D., & Shepherd, J. M. (2008). Global precipitation measurement. In *Precipitation: Advances in measurement, estimation and prediction*, 131-169. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-77655-0_6

Hu, Q., Jiang, D., Lang, X., & Xu, B. (2018). Moisture sources of the Chinese Loess Plateau during 1979–2009. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *509*, 156–163. https://doi.org/10.1016/j.palaeo.2016.12.030

Hu, X., Chu, L., Pei, J., Liu, W., & Bian, J. (2021). Model complexity of deep learning: A survey. *Knowledge and Information Systems*, *63*, 2585-2619. https://doi.org/10.1007/s10115-021-01605-0

Huang, Z., Tan, X., & Liu, B. (2024). Relative contributions of large-scale atmospheric circulation dynamics and anthropogenic warming to the unprecedented 2022 Yangtze River Basin heatwave. *Journal of Geophysical Research: Atmospheres*, *129*(4), e2023JD039330.

https://doi.org/10.1029/2023JD039330

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., & Wolff, D. B. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, *8*(1), 38–55. https://doi.org/10.1175/JHM560.1

Hüllermeier, E., & Waegeman, W. (2021). Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods. *Machine learning*, *110*(3), 457-506.

Humphrey, V., Berg, A., Ciais, P., Gentine, P., Jung, M., Reichstein, M., Seneviratne, S. I., & Frankenberg, C. (2021). Soil moisture–atmosphere feedback dominates land carbon uptake variability. *Nature*, *592*(7852), 65–69. https://doi.org/10.1038/s41586-021-03325-5

Imbens, G. W., & Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge university press.

Insua-Costa, D., & Miguez-Macho, G. (2018). A new moisture tagging capability in the Weather Research and Forecasting model: Formulation, validation and application to the 2014 Great Lake-effect snowstorm. *Earth System Dynamics*, *9*(1), 167–185. https://doi.org/10.5194/esd-9-167-2018

International Federation of Red Cross Red Crescent Societies (2020). Viet Nam: Drought and saltwater intrusion. Final Report.

Jabbar, H., & Khan, R. Z. (2015). Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study). *Computer Science, Communication and Instrumentation Devices*, *70*(10.3850), 978-981.

Jadhav, B. D., Marotrao Ghate, P., Kota, P. N., Chavan, S. D., & Chopade, P. B. (2024). An optimized network for drought prediction using satellite images. *Remote Sensing Applications: Society and Environment*, *36*(June), 101278. https://doi.org/10.1016/j.rsase.2024.101278

Janssen, R. H. H., Vilà-Guerau De Arellano, J., Jimenez, J. L., Ganzeveld, L. N., Robinson, N. H., Allan, J. D., Coe, H., & Pugh, T. A. M. (2013). Influence of boundary layer dynamics and isoprene chemistry on the organic aerosol budget in a tropical forest. *Journal of Geophysical Research Atmospheres*, *118*(16), 9351–9366. https://doi.org/10.1002/jgrd.50672

Javed, T., Li, Y., Rashid, S., Li, F., Hu, Q., Feng, H., Chen, X., Ahmad, S., Liu, F., & Pulatov, B. (2021). Performance and relationship of four different agricultural drought indices for drought monitoring in China's mainland using remote sensing data. *Science of the Total Environment*, *759*. https://doi.org/10.1016/j.scitotenv.2020.143530

Javier, P. J. E. (2021). causal-ccm a Python implementation of Convergent Cross Mapping: Version 0.3.3 [Software]. *GitHub*. https://github.com/PrinceJavier/causal_ccm

Jha, S., Das, J., Sharma, A., Hazra, B., & Goyal, M. K. (2019). Probabilistic evaluation of vegetation drought likelihood and its implications to resilience across India. *Global and Planetary Change*, *176*(July 2018), 23–35. https://doi.org/10.1016/j.gloplacha.2019.01.014

Jiang, S., Ren, L., Hong, Y., Yong, B., Yang, X., Yuan, F., & Ma, M. (2012). Comprehensive evaluation of multi-satellite precipitation products with a dense rain gauge network and optimally merging their simulated hydrological flows using the Bayesian model averaging method. *Journal of Hydrology*, *452–453*, 213–225. https://doi.org/10.1016/j.jhydrol.2012.05.055

Jiang, S., Wei, L., Ren, L., Xu, C. Y., Zhong, F., Wang, M., Zhang, L., Yuan, F., & Liu, Y. (2021). Utility of integrated IMERG precipitation and GLEAM potential evapotranspiration products for drought monitoring over mainland China. *Atmospheric Research*, *247*(May 2020), 105141. https://doi.org/10.1016/j.atmosres.2020.105141

Jiang, W., & Luo, J. (2022). An evaluation of machine learning and deep learning models for drought prediction using weather data. *Journal of Intelligent and Fuzzy Systems*, *43*(3), 3611–3626. https://doi.org/10.3233/JIFS-212748

John, A., Douville, H., Ribes, A., & Yiou, P. (2022). Quantifying CMIP6 model uncertainties in extreme precipitation projections. *Weather and Climate Extremes*, *36*, 100435.

John, J. G., Blanton, C., McHugh, C., Radhakrishnan, A., Rand, K., Vahlenkamp, H., Wilson, C., Zadeh, N. T., Dunne, J. P., Dussin, R., Horowitz, L. W., Krasting, J. P., Lin, P., Malyshev, S., Naik, V., Ploshay, J., Shevliakova, E., Silvers, L., Stock, C.,Winton, M., Zeng, Y. (2018). NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 ScenarioMIP ssp585 [Dataset]. *Earth System Grid Federation*. https://doi.org/10.22033/ESGF/CMIP6.8706

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G., Keeley, S. P. E., Mogensen, K., Zuo, H., & Monge-Sanz, B. M. (2019). SEAS5: The new ECMWF seasonal forecast system. *Geoscientific Model Development*, *12*(3), 1087–1117. https://doi.org/10.5194/gmd-12-1087-2019

Joussaume, S., Sadourny, R., & Jouzel, J. (1984). A general circulation model of water isotope cycles in the atmosphere. *Nature*, *311*(5981), 24-29. https://doi.org/10.1038/311024a0

Jouzel, J., Delaygue, G., Landais, A., Masson-Delmotte, V., Risi, C., & Vimeux, F. (2013). Water isotopes as tools to document oceanic sources of precipitation. *Water Resources Research*, *49*(11), 7469–7486. https://doi.org/10.1002/2013WR013508

Joyce, R. J., Janowiak, J. E., Arkin, P. A., & Xie, P. (2004). CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, *5*(3), 487–503. https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2

Justice, C. O., Townshend, J. R. G., Vermote, E. F., Masuoka, E., Wolfe, R. E., Saleous, N., Roy, D. P., & Morisette, J. T. (2002). An overview of MODIS Land data processing and product status. *Remote Sensing of Environment*, *83*(1–2), 3–15. https://doi.org/10.1016/S0034-4257(02)00084-6

Kafle, S. K. (2017). Disaster early warning systems in Nepal: Institutional and operational frameworks. *Journal of Geography & Natural Disasters*, *7*(2), 2167-0587. https://doi.org/10.4172/2167-0587.1000196

Kalisch, M., & Bühlmann, P. (2007). Estimating high-dimensional directed acyclic graphs with the PC-algorithm. *Journal of Machine Learning Research*, *8*, 613–636.

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, *147*(July 2017), 70–90. https://doi.org/10.1016/j.compag.2018.02.016

Kang, H., & Sridhar, V. (2021). A near-term drought assessment using hydrological and climate forecasting in the Mekong River Basin. *International Journal of Climatology*, *41*(S1), E2497-E2516. https://doi.org/10.1002/joc.6860

Kang, H., Sridhar, V., & Ali, S. A. (2022). Climate change impacts on conventional and flash droughts in the Mekong River Basin. *Science of the Total Environment*, *838*(April), 155845. https://doi.org/10.1016/j.scitotenv.2022.155845

Kang, H., Sridhar, V., Mainuddin, M., & Trung, L. D. (2021). Future rice farming threatened by drought in the Lower Mekong Basin. *Scientific Reports*, *11*(1), 1–15. https://doi.org/10.1038/s41598-021-88405-2

Kantoush, S., Binh, D. Van, Sumi, T., & Trung, L. V. (2017). Impact of upstream hydropower dams and climate change on hydrodynamics of Vietnamese Mekong Delta. *Journal of Japan Society of Civil Engineers, Ser. B1 (Hydraulic Engineering), 73*(4), I_109-I_114. https://doi.org/10.2208/jscejhe.73.I_109

Karl, T. R. (1986). The sensitivity of the Palmer Drought Severity Index and Palmer's Z-index to their calibration coefficients including potential evapotranspiration. *Journal of Climate and Applied Meteorology*, 77-86. http://www.jstor.org/stable/26182460.

Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H. A., & Kumar, V. (2018). Machine learning for the geosciences: Challenges and opportunities. *IEEE Transactions on Knowledge and Data Engineering*, *31*(8), 1544-1554. https://doi.org/10.1109/TKDE.2018.2861006

Katiraie-Boroujerdy, P. S., Nasrollahi, N., Hsu, K. lin, & Sorooshian, S. (2016). Quantifying the reliability of four global datasets for drought monitoring over a semiarid region. *Theoretical and Applied Climatology*, *123*(1–2), 387–398. https://doi.org/10.1007/s00704-014-1360-3

Kaur, A., & Sood, S. K. (2020). Deep learning based drought assessment and prediction framework. *Ecological Informatics, 57*(March), 101067. https://doi.org/10.1016/j.ecoinf.2020.101067

Keovilignavong, O., Nguyen, T. H., & Hirsch, P. (2021). Reviewing the causes of Mekong drought before and during 2019–20. *International Journal of Water Resources Development*, *00*(00), 1-21. https://doi.org/10.1080/07900627.2021.1967112

Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M. J., Font, J., Reul, N., Gruhier, C., Juglea, S. E., Drinkwater, M. R., Hahne, A., Martin-Neira, M., & Mecklenburg, S. (2010). The SMOS L: New tool for monitoring key elements ofthe global water cycle. *Proceedings of the IEEE, 98*(5), 666–687. https://doi.org/10.1109/JPROC.2010.2043032

Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Martinuzzi, J. M., Font, J., & Berger, M. (2001). Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(8), 1729–1735. https://doi.org/10.1109/36.942551

Keyantash, J. A., & Dracup, J. A. (2004). An aggregate drought index: Assessing drought severity based on fluctuations in the hydrologic cycle and surface water storage. *Water Resources Research*, *40*(9), 1–14. https://doi.org/10.1029/2003WR002610

Keyantash, J., & Dracup, J. A. (2002). The quantification of drought: an evaluation of drought indices. *Bulletin of the American Meteorological Society*, *83*(8), 1167-1180. https://doi.org/10.1175/1520-0477-83.8.1167

Keys, P. W., Barnes, E. A., Van Der Ent, R. J., & Gordon, L. J. (2014). Variability of moisture recycling using a precipitationshed framework. *Hydrology and Earth System Sciences*, *18*(10), 3937–3950. https://doi.org/10.5194/hess-18-3937-2014

Keys, P. W., Van Der Ent, R. J., Gordon, L. J., Hoff, H., Nikoli, R., & Savenije, H. H. G. (2012). Analyzing precipitationsheds to understand the vulnerability of rainfall dependent regions. *Biogeosciences*, *9*(2), 733–746. https://doi.org/10.5194/bg-9-733-2012

Khan, H., Vasilescu, L. G., & Khan, A. (2008). Disaster management cycle-a theoretical approach. *Journal of Management and Marketing*, *6*(1), 43-50.

Khan, M. I., & Maity, R. (2024). Development of a long-range hydrological drought prediction framework using deep learning. *Water Resources Management*, *38*(4), 1497–1509. https://doi.org/10.1007/s11269-024-03735-w

Khan, M. M. H., Muhammad, N. S., & El-Shafie, A. (2020). Wavelet based hybrid ANN-ARIMA models for meteorological drought forecasting. *Journal of Hydrology*, *590*(July). https://doi.org/10.1016/j.jhydrol.2020.125380

Khan, N., Sachindra, D. A., Shahid, S., Ahmed, K., Shiru, M. S., & Nawaz, N. (2020). Prediction of droughts over Pakistan using machine learning algorithms. *Advances in Water Resources*, *139*(August 2019), 103562. https://doi.org/10.1016/j.advwatres.2020.103562

Kidd, C. (2001). Satellite rainfall climatology: A review. *International Journal of Climatology*, *21*(9), 1041–1066. https://doi.org/10.1002/joc.635

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., & Kirschbaum, D. B. (2017). So, how much of the Earth's surface is covered by rain gauges? *Bulletin of the American Meteorological Society*, *98*(1), 69–78. https://doi.org/10.1175/BAMS-D-14-00283.1

Kim, J. J., & Guha-Sapir, D. (2012). Famines in Africa: is early warning early enough?. *Global health action*, *5*(1), 18481. https://doi.org/10.3402/gha.v5i0.18481

Kingston, D. G., Stagge, J. H., Tallaksen, L. M., & Hannah, D. M. (2015). European-scale drought: Understanding connections between atmospheric circulation and meteorological drought indices. *Journal of Climate*, *28*(2), 505–516. https://doi.org/10.1175/JCLI-D-14-

00001.1

Kingston, D. G., Thompson, J. R., & Kite, G. (2011). Uncertainty in climate change projections of discharge for the Mekong River Basin. *Hydrology and Earth System Sciences*, *15*(5), 1459–1471. https://doi.org/10.5194/hess-15-1459-2011

Kirtman, B. P., Min, D., Infanti, J. M., Kinter, J. L., III, Paolino, D. A., Zhang, Q., van den Dool, H., Saha, S., Mendez, M. P., Becker, E., Peng, P., Tripp, P., Huang, J., DeWitt, D. G., Tippett, M. K., Barnston, A. G., Li, S., Rosati, A., Schubert, S. D., Rienecker, M., Suarez, M., Li, Z. E., Marshak, J., Lim, Y., Tribbia, J., Pegion, K., Merryfield, W. J., Denis, B., & Wood, E. F. (2014). The North American Multimodel Ensemble: Phase-1 seasonal-to-interannual prediction; Phase-2 toward developing intraseasonal prediction. *Bulletin of the American Meteorological Society*, *95*(4), 585-601. https://doi.org/10.1175/BAMS-D-12-00050.1

Knippertz, P., Wernli, H., & Gläser, G. (2013). A global climatology of tropical moisture exports. *Journal of Climate*, *26*(10), 3031–3045. https://doi.org/10.1175/JCLI-D-12-00401.1

Kogan, F. (2002). World droughts in the new millennium from AVHRR-based vegetation health indices. *Eos, Transactions American Geophysical Union*, *83*(48), 557-563. https://doi.org/10.1029/2002EO000382

Kogan, F. N. (1995). Application of vegetation index and brightness temperature for drought detection. *Advances in space research*, *15*(11), 91-100. https://doi.org/10.1016/0273-1177(95)00079-T

Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C. T., Kanae, S., Kowalczyk, E., Lawrence, D., Liu, P., Lu, C. H., Malyshev, S., McAvaney, B., Mitchell, K., Mocko, D., Oki, T., Oleson, K., Pitman, A., Sud, Y. C., Taylor, C. M., Verseghy, D., Vasic, R., Xue, Y., & Yamada, T. (2004). Regions of strong coupling between soil moisture and precipitation. *Science*, *305*(5687), 1138-1140. https://doi.org/10.1126/science.1100217

Koster, R. D., Sud, Y. C. , Guo, Z., Dirmeyer, P. A., Bonan, G., Oleson, K. W., Chan, E., Verseghy, D., Cox, P., Davies, H., Kowalczyk, E., Gordon, C. T. , Kanae, S., Lawrence, D., Liu, P., Mocko, D., Lu, C., Mitchell, K., Malyshev, S., McAvaney, B., Oki, T., Yamada, T., Pitman, A., Taylor, C. M., Vasic, R., & Xue, Y. (2006). GLACE: The Global Land–Atmosphere Coupling Experiment. Part I: Overview. *Journal of Hydrometeorology*, *7*(4), 590-610. https://doi.org/10.1175/JHM510.1

Koster, R., Jouzel, J., Suozzo, R., Russell, G., Broecker, W., Rind, D., & Eagleson, P. (1986).

Global sources of local precipitation as determined by the NASA/GISS GCM. *Geophysical research letters*, *13*(2), 121-124. https://doi.org/10.1029/GL013i002p00121

Krich, C., Runge, J., Miralles, D. G., Migliavacca, M., Perez-Priego, O., El-Madany, T., Carrara, A., & Mahecha, M. D. (2020). Estimating causal networks in biosphere–atmosphere interaction with the PCMCI approach. *Biogeosciences*, *17*, 1033-1061. https://doi.org/10.5194/bg-17-1033-2020.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, *25*.

Kumar, N., Poonia, V., Gupta, B. B., & Goyal, M. K. (2021). A novel framework for risk assessment and resilience of critical infrastructure towards climate change. *Technological Forecasting and Social Change*, *165*, 120532. https://doi.org/10.1016/j.techfore.2020.120532

Kuwayama, Y., Thompson, A., Bernknopf, R., Zaitchik, B., & Vail, P. (2019). Estimating the impact of drought on agriculture using the U.S. Drought Monitor. *American Journal of Agricultural Economics*, *101*(1), 193–210. https://doi.org/10.1093/ajae/aay037

Lai, G., Chang, W. C., Yang, Y., & Liu, H. (2018). Modeling long- and short-term temporal patterns with deep neural networks. *41st International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2018*, (July), 95–104. https://doi.org/10.1145/3209978.3210006

Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. *Advances in neural information processing systems*, *30*.

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., Ravuri, S., Ewalds, T., Eaton-Rosen, Z., Hu, W., Merose, A., Hoyer, S., Holland, G., Vinyals, O., Stott, J., Pritzel, A., Mohamed, S., & Battaglia, P. (2023). Learning skillful medium-range global weather forecasting. *Science*, *382*(6677), 1416–1422. https://doi.org/10.1126/science.adi2336

Lansu, E. M., van Heerwaarden, C. C., Stegehuis, A. I., & Teuling, A. J. (2020). Atmospheric aridity and apparent soil moisture drought in European forest during heat waves. *Geophysical Research Letters*, *47*(6). https://doi.org/10.1029/2020GL087091

Latifoğlu, L., & Özger, M. (2023). A novel approach for high-performance estimation of SPI data in drought prediction. *Sustainability*, *15*(19), 14046.

https://doi.org/10.3390/su151914046

Lavane, K., Kumar, P., Meraj, G., Han, T. G., Ngan, L. H. B., Lien, B. T. B., Van Ty, T., Thanh, N. T., Downes, N. K., Nam, N. D. G., Minh, H. V. T., Singh, S. K., & Kanga, S. (2023). Assessing the effects of drought on rice yields in the Mekong Delta. *Climate*, *11*(1). https://doi.org/10.3390/cli11010013

Le, J. A., El-Askary, H. M., Allali, M., & Struppa, D. C. (2017). Application of recurrent neural networks for drought projections in California. *Atmospheric Research*, *188*, 100–106. https://doi.org/10.1016/j.atmosres.2017.01.002

Le, M. H., Kim, H., Moon, H., Zhang, R., Lakshmi, V., & Nguyen, L. B. (2020). Assessment of drought conditions over Vietnam using standardized precipitation evapotranspiration index, MERRA-2 re-analysis, and dynamic land cover. *Journal of Hydrology: Regional Studies*, *32*, 100767. https://doi.org/10.1016/j.ejrh.2020.100767

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278-2324. https://doi.org/10.1109/5.726791

Lee, S. K., & Dang, T. A. (2018). Evaluating drought events under influence of El-nino phenomenon: A case study of mekong delta area, Vietnam. *Journal of Agrometeorology*, *20*(4), 275–279.

Lee, S. K., & Dang, T. A. (2019). Comparison of drought indices (SPI and RDI) over Mekong Delta area of Vietnam. *Journal of Agrometeorology*, *21*(2), 213-216. https://doi.org/10.54386/jam.v21i2.236

Lee, T., Singh, V. P., & Cho, K. H. (2021). *Deep learning for hydrometeorology and environmental science* (pp. 21-25). Springer.

Li, J., Wang, Z., Wu, X., Xu, C. Y., Guo, S., Chen, X., & Zhang, Z. (2021a). Robust meteorological drought prediction using antecedent SST fluctuations and machine learning. *Water Resources Research*, *57*(8), 1–20. https://doi.org/10.1029/2020WR029413

Li, J., Wang, Z., Wu, X., Zscheischler, J., Guo, S., & Chen, X. (2021b). A standardized index for assessing sub-monthly compound dry and hot conditions with application in China. *Hydrology and Earth System Sciences*, *25*(3), 1587–1601. https://doi.org/10.5194/hess-25-

1587-2021

Li, L., Li, W., & Barros, A. P. (2013). Atmospheric moisture budget and its regulation of the summer precipitation variability over the Southeastern United States. *Climate Dynamics*, *41*(3–4), 613–631. https://doi.org/10.1007/s00382-013-1697-9

Li, Q., Ye, A., Zhang, Y., & Zhou, J. (2022). The peer-to-peer type propagation from meteorological drought to soil moisture drought occurs in areas with strong land-atmosphere interaction. *Water Resources Research*, *58*(9). https://doi.org/10.1029/2022WR032846

Li, Y., Lu, H., Yang, K., Wang, W., Tang, Q., Khem, S., Yang, F., & Huang,Y. (2021). Meteorological and hydrological droughts in Mekong River Basin and surrounding areas under climate change. *Journal of Hydrology: Regional Studies*, *36*(June), 100873. https://doi.org/10.1016/j.ejrh.2021.100873

Li, Y., Su, F., Tang, Q., Gao, H., Yan, D., Peng, H., & Xiao, S. (2022). Contributions of moisture sources to precipitation in the major drainage basins in the Tibetan Plateau. *Science China Earth Sciences*, *65*(6), 1088–1103. https://doi.org/10.1007/s11430-021-9890-6

Li, Z., Hao, Z., Shi, X., Déry, S. J., Li, J., Chen, S. & Li, Y. (2016) An agricultural drought index to incorporate the irrigation process and reservoir operations: a case study in the Tarim River Basin. *Global and Planetary Change*, *143*, 10-20. https://doi.org/10.1016/j.gloplacha.2016.05.008.

Lin, Q., Wu, Z., Zhang, Y., Peng, T., Chang, W., & Guo, J. (2023). Propagation from meteorological to hydrological drought and its application to drought prediction in the Xijiang River basin, South China. *Journal of Hydrology*, *617*(PB), 128889. https://doi.org/10.1016/j.jhydrol.2022.128889

Liu, B., Tan, X., Gan, T. Y., Chen, X., Lin, K., Lu, M., & Liu, Z. (2020). Global atmospheric moisture transport associated with precipitation extremes: Mechanisms and climate change impacts. *Wiley Interdisciplinary Reviews: Water*, *7*(2), 1–25. https://doi.org/10.1002/wat2.1412

Liu, J., Rahmani, F., Lawson, K., & Shen, C. (2022b). A multiscale deep learning model for soil moisture integrating satellite and in situ data. *Geophysical Research Letters*, *49*(7). https://doi.org/10.1029/2021GL096847

Liu, J., Yuan, X., Zeng, J., Jiao, Y., Li, Y., Zhong, L., & Yao, L. (2022a). Ensemble streamflow forecasting over a cascade reservoir catchment with integrated hydrometeorological modeling and machine learning. *Hydrology and Earth System Sciences*, *26*(2), 265–278. https://doi.org/10.5194/hess-26-265-2022

Liu, L., Gu, H., Xie, J., & Xu, Y. P. (2021). How well do the ERA-Interim, ERA-5, GLDAS-2.1 and NCEP-R2 reanalysis datasets represent daily air temperature over the Tibetan Plateau? *International Journal of Climatology*, *41*(2), 1484–1505. https://doi.org/10.1002/joc.6867

Liu, W. T., & Kogan, F. N. (1996). Monitoring regional drought using the vegetation condition index. *International Journal of Remote Sensing*, *17*(14), 2761-2782. https://doi.org/10.1080/01431169608949106

Liu, W., Huang, Y., & Wang, H. (2024). Effective deep learning seasonal prediction model for summer drought over China. *Earth's Future*, *12*(3), e2023EF004409. https://doi.org/10.1029/2023EF004409

Liu, X., Zhu, X., Pan, Y., Li, S., Liu, Y., & Ma, Y. (2016). Agricultural drought monitoring: Progress, challenges, and prospects. *Journal of Geographical Sciences*, *26*(6), 750–767. https://doi.org/10.1007/s11442-016-1297-9

Liu, X., Zhu, X., Zhang, Q., Yang, T., Pan, Y., & Sun, P. (2020). A remote sensing and artificial neural network-based integrated agricultural drought index: Index development and applications. *Catena*, *186*(December 2019), 104394. https://doi.org/10.1016/j.catena.2019.104394

Liu, Y., & Wu, L. (2016). Geological disaster recognition on optical remote sensing images using deep learning. *Procedia Computer Science*, *91*(Itqm), 566–575. https://doi.org/10.1016/j.procs.2016.07.144

Liu, Y., Liu, Y., & Wang, W. (2019). Inter-comparison of satellite-retrieved and Global Land Data Assimilation System-simulated soil moisture datasets for global drought analysis. *Remote Sensing of Environment*, *220*(February 2018), 1–18. https://doi.org/10.1016/j.rse.2018.10.026

Liu, Y., Racah, E., Prabhat, Correa, J., Khosrowshahi, A., Lavers, D., Kunkel, K., Wehner, M., & Collins, W. (2016). *Application of deep convolutional neural networks for detecting extreme weather in climate datasets*. https://doi.org/10.475/123

Liu, Y., Zheng, H., Feng, X., & Chen, Z. (2017). Short-term traffic flow prediction with Conv-LSTM. In *2017 9th international conference on wireless communications and signal processing (WCSP)* (pp. 1-6). IEEE. https://doi.org/10.1109/WCSP.2017.8171119

Lloyd-Hughes, B. (2014). The impracticality of a universal drought definition. *Theoretical and Applied Climatology*, *117*(3–4), 607–611. https://doi.org/10.1007/s00704-013-1025-7

Loc, H. H., Low Lixian, M., Park, E., Dung, T. D., Shrestha, S., & Yoon, Y. J. (2021). How the saline water intrusion has reshaped the agricultural landscape of the Vietnamese Mekong Delta, a review. *Science of the Total Environment*, *794*, 148651. https://doi.org/10.1016/j.scitotenv.2021.148651

Loquercio, A., Segu, M., & Scaramuzza, D. (2020). A general framework for uncertainty estimation in deep learning. *IEEE Robotics and Automation Letters*, *5*(2), 3153-3160.

Lorenz, R., Argüeso, D., Donat, M. G., Pitman, A. J., Hurk, B. Van Den, Berg, A., Lawrence, D. M., Chéruy, F., Ducharne, A., Hagemann, S., Meier, A., Milly, P. C. D., & Seneviratne, S. I. (2016). Influence of land-atmosphere feedbacks on temperature and precipitation extremes in the GLACE-CMIP5 ensemble. *Journal of Geophysical Research*, *121*(2), 607–623. https://doi.org/10.1002/2015JD024053

Loukas, A., Vasiliades, L., & Tzabiras, J. (2008). Climate change effects on drought severity, Adv. Geosci., 17, 23–29, https://doi.org/10.5194/adgeo-17-23-2008

Lovato, T., Peano, D., Butenschön, M. (2021). CMCC CMCC-ESM2 model output prepared for CMIP6 ScenarioMIP ssp585 [Dataset]. *Earth System Grid Federation*. https://doi.org/10.22033/ESGF/CMIP6.13259

Lu, J., Carbone, G. J., & Grego, J. M. (2019). Uncertainty and hotspots in 21st century projections of agricultural drought from CMIP5 models. *Scientific Reports*, *9*(1), 1–12. https://doi.org/10.1038/s41598-019-41196-z

Lu, J., Carbone, G. J., Huang, X., Lackstrom, K., & Gao, P. (2020). Mapping the sensitivity of agriculture to drought and estimating the effect of irrigation in the United States, 1950–2016. *Agricultural and Forest Meteorology*, *292–293*(May), 108124. https://doi.org/10.1016/j.agrformet.2020.108124

Lu, J., Liu, A., Dong, F., Gu, F., Gama, J., & Zhang, G. (2018). Learning under concept drift: A review. *IEEE transactions on knowledge and data engineering*, *31*(12), 2346-2363. https://doi.org/10.1109/TKDE.2018.2876857

Lu, X. X., Li, S., Kummu, M., Padawangi, R., & Wang, J. J. (2014). Observed changes in the water flow at Chiang Saen in the lower Mekong: Impacts of Chinese dams? *Quaternary International*, *336*, 145-157. https://doi.org/10.1016/j.quaint.2014.02.006

Luca, A. R., Ursuleanu, T. F., Gheorghe, L., Grigorovici, R., Iancu, S., Hlusneac, M., & Grigorovici, A. (2022). Impact of quality, type and volume of data used by deep learning models in the analysis of medical images. *Informatics in Medicine Unlocked*, *29*, 100911. https://doi.org/10.1016/j.imu.2022.100911

Ma, F., Ye, A., Deng, X., Zhou, Z., Liu, X., Duan, Q., Xu, J., Miao, C., Di, Z., & Gong. W. (2016). Evaluating the skill of NMME seasonal precipitation ensemble predictions for 17 hydroclimatic regions in continental China. *International Journal of Climatology*, *36*(1), 132–144. https://doi.org/10.1002/joc.4333

MacAllister, D. J., MacDonald, A. M., Kebede, S., Godfrey, S., & Calow, R. (2020). Comparative performance of rural water supplies during drought. *Nature Communications*, *11*(1), 1–13. https://doi.org/10.1038/s41467-020-14839-3

Madadgar, S., AghaKouchak, A., Farahmand, A., & Davis, S. J. (2017). Probabilistic estimates of drought impacts on agricultural production. *Geophysical Research Letters*, *44*(15), 7799–7807. https://doi.org/10.1002/2017GL073606

Madadgar, S., AghaKouchak, A., Shukla, S., Wood, A. W., Cheng, L., Hsu, K. L., & Svoboda, M. (2016). A hybrid statistical-dynamical framework for meteorological drought prediction: Application to the southwestern United States. *Water Resources Research*, *52*(7), 5095-5110. https://doi.org/10.1002/2015WR018547

Malik, A., & Kumar, A. (2020). Meteorological drought prediction using heuristic approaches based on effective drought index: a case study in Uttarakhand. *Arabian Journal of Geosciences*, *13*(6). https://doi.org/10.1007/s12517-020-5239-6

Malik, A., Tikhamarine, Y., Souag-Gamane, D., Rai, P., Sammen, S. S., & Kisi, O. (2021). Support vector regression integrated with novel meta-heuristic algorithms for meteorological drought prediction. *Meteorology and Atmospheric Physics*, *133*(3), 891–909. https://doi.org/10.1007/s00703-021-00787-0

Mann, M. E., & Gleick, P. H. (2015). Climate change and California drought in the 21st century. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(13), 3858–3859. https://doi.org/10.1073/pnas.1503667112

Maranghi, M., Anagnostopoulos, A., Cannistraci, I., Chatzigiannakis, I., Croce, F., Di Teodoro, G., Gentile, M., Grani, G., Lenzerini, M., Leonardi, S., Mastropietro, A., Palagi, L., Pappa, M., Rosati, R., Valentini R., & Velardi, P. (2022). AI-based Data Preparation and Data

Analytics in Healthcare: The Case of Diabetes. *arXiv preprint arXiv:2206.06182*. https://doi.org/10.48550/arXiv.2206.06182

Markatou, M., Tian, H., Biswas, S., & Hripcsak, G. (2005). Analysis of variance of cross-validation estimators of the generalization error. *Journal of Machine Learning Research*, *6*, 1127–1168.

Márquez-Grajales, A., Villegas-Vega, R., Salas-Martínez, F., Acosta-Mesa, H. G., & Mezura-Montes, E. (2024). Characterizing drought prediction with deep learning: A literature review. *MethodsX*, 102800. https://doi.org/10.1016/j.mex.2024.102800

Mårtensson, G., Ferreira, D., Granberg, T., Cavallin, L., Oppedal, K., Padovani, A., Rektorova, I., Bonanni, L., Pardini, M., Kramberger, M. G., Taylor, J. P., Hort, J., Snædal, J., Kulisevsky, J., Blanc, F., Antonini, A., Mecocci, P., Vellas, B., Tsolaki, M., Kłoszewska, I., Soininen, H., Lovestone, S., Simmons, A., Aarsland, D., & Westman, E. (2020). The reliability of a deep learning model in clinical out-of-distribution MRI data: a multicohort study. *Medical Image Analysis*, *66*, 101714. https://doi.org/10.1016/j.media.2020.101714

Martínez-Fernández, J., González-Zamora, A., Sánchez, N., & Gumuzzio, A. (2015). A soil water based index as a suitable agricultural drought indicator. *Journal of Hydrology*, *522*, 265–273. https://doi.org/10.1016/j.jhydrol.2014.12.051

Masegosa, A. R., Martínez, A. M., Ramos-López, D., Langseth, H., Nielsen, T. D., & Salmerón, A. (2020). Analyzing concept drift: A case study in the financial sector. *Intelligent Data Analysis*, *24*(3), 665-688. https://doi.org/10.3233/IDA-194515

Massmann, A., Gentine, P., & Runge, J. (2021). *Causal inference for process understanding in Earth sciences*. 1–24. http://arxiv.org/abs/2105.00912

Masupha, T. E., Moeletsi, M. E., & Tsubo, M. (2021). Prospects of an agricultural drought early warning system in South Africa. *International Journal of Disaster Risk Reduction*, *66*, 102615. https://doi.org/10.1016/j.ijdrr.2021.102615

Masupha, T. E., Moeletsi, M. E., & Tsubo, M. (2024). A transformative framework reshaping sustainable drought risk management through advanced early warning systems. *Iscience*, *27*(7). https://doi.org/10.1016/j.isci.2024.110066

McKee, T. B., Doesken, N. J., & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology* (Vol. 17, No. 22, pp. 179-183).

McKee, T. B., Doesken, N. J., & Kleist, J. (1995). Drought monitoring with multiple time scales. PROCEEDINGS OF THE CONFERENCE ON APPLIED CLIMATOLOGY.

Mehran, A., Mazdiyasni, O., & AghaKouchak, A. (2015). A hybrid framework for assessing socioeconomic drought: Linking climate variability, local resilience, and demand. *Journal of Geophysical Research: Atmospheres*, *120*(15), 7520-7533. https://doi.org/10.1038/175238c0

Menéndez, C. G., Zaninelli, P. G., Carril, A. F., & Sánchez, E. (2016). Hydrological cycle, temperature, and land surface atmosphere interaction in the La Plata Basin during summer: response to climate change. *Climate Research*, *68*(2-3), 231-241. https://doi.org/10.3354/cr01373

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., & Houston, T. G. (2012). An overview of the global historical climatology network-daily database. *Journal of Atmospheric and Oceanic Technology, 29*(7), 897–910. https://doi.org/10.1002/2015JD023147

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, *26*.

Min, S., Lee, B., & Yoon, S. (2017). Deep learning in bioinformatics. *Briefings in bioinformatics*, *18*(5), 851-869. https://doi.org/10.1093/bib/bbw068

Miralles, D. G., Gentine, P., Seneviratne, S. I., & Teuling, A. J. (2019). Land-atmospheric feedbacks during droughts and heatwaves: state of the science and current challenges. *Annals of the New York Academy of Sciences*, *1436*(1), 19-35. https://doi.org/10.1111/nyas.13912

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., & Dolman, A. J. (2011). Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences, 15*(2), 453–469. https://doi.org/10.5194/hess-15-453-2011

Miralles, D. G., Teuling, A. J., Van Heerwaarden, C. C., & De Arellano, J. V. G. (2014). Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nature Geoscience, 7*(5), 345–349. https://doi.org/10.1038/ngeo2141

Mishra, A. K., & Singh, V. P. (2010). A review of drought concepts. *Journal of Hydrology*, *391*(1–2), 202–216. https://doi.org/10.1016/j.jhydrol.2010.07.012

Mishra, A., Vu, T., Veettil, A. V., & Entekhabi, D. (2017). Drought monitoring with soil moisture active passive (SMAP) measurements. *Journal of Hydrology*, *552*(January 2015), 620–632. https://doi.org/10.1016/j.jhydrol.2017.07.033

Miyan, M. A. (2015). Droughts in asian least developed countries: Vulnerability and sustainability. *Weather and Climate Extremes*, *7*, 8–23. https://doi.org/10.1016/j.wace.2014.06.003

Mizoguchi, Y., Miyata, A., Ohtani, Y., Hirata, R., & Yuta, S. (2009). A review of tower flux observation sites in Asia. *Journal of Forest Research*, *14*(1), 1–9. https://doi.org/10.1007/s10310-008-0101-9

Mizuta, R., Yoshimura, H., Ose, T., Hosaka, M., Yukimoto, S. (2019). MRI MRI-AGCM3-2-H model output prepared for CMIP6 HighResMIP highresSST-future [Dataset]. *Earth System Grid Federation*. https://doi.org/10.22033/ESGF/CMIP6.10972

Mo, K. C., & Lettenmaier, D. P. (2020). Prediction of flash droughts over the United States. *Journal of Hydrometeorology*, *21*(8), 1793–1810. https://doi.org/10.1175/JHM-D-19-0221.1

Mo, K. C., & Lyon, B. (2015). Global meteorological drought prediction using the North American multi-model ensemble. *Journal of Hydrometeorology*, *16*(3), 1409–1424. https://doi.org/10.1175/JHM-D-14-0192.1

Mohammed, S., Elbeltagi, A., Bashir, B., Alsafadi, K., Alsilibe, F., Alsalman, A., Zeraatpisheh, M., Széles, A., & Harsányi, E. (2022). A comparative analysis of data mining techniques for agricultural and hydrological drought prediction in the eastern Mediterranean. *Computers and Electronics in Agriculture*, *197*(March). https://doi.org/10.1016/j.compag.2022.106925

Mohd Noor, M. H., Tan, S. Y., & Ab Wahab, M. N. (2022). Deep temporal Conv-LSTM for activity recognition. *Neural Processing Letters*, *54*(5), 4027-4049. https://doi.org/10.1007/s11063-022-10799-5

Mol, A., Tait, A., & Macara, G. (2017). An automated drought monitoring system for New Zealand. *Weather and Climate*, *37*(1), 23-36. https://doi.org/10.2307/26735444

Mueller, B., & Seneviratne, S. I. (2012). Hot days induced by precipitation deficits at the global scale. *Proceedings of the national academy of sciences*, *109*(31), 12398-12403. https://doi.org/10.1073/pnas.1204330109

Mukherjee, S., Mishra, A., & Trenberth, K. E. (2018). Climate change and drought: A perspective on drought indices. *Current Climate Change Reports*, *4*(2), 145–163. https://doi.org/10.1007/s40641-018-0098-x

Müller, L. M., & Bahn, M. (2022). Drought legacies and ecosystem responses to subsequent drought. *Global Change Biology*, *28*(17), 5086–5103. https://doi.org/10.1111/gcb.16270

Munappy, A. R., Bosch, J., Olsson, H. H., Arpteg, A., & Brinne, B. (2019). Data management challenges for deep learning. In *2019 45th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)* (pp. 140-147). IEEE. https://doi.org/10.1109/SEAA.2019.00030

Munappy, A. R., Bosch, J., Olsson, H. H., Arpteg, A., & Brinne, B. (2022). Data management for production quality deep learning models: Challenges and solutions. *Journal of Systems and Software*, *191*, 111359. https://doi.org/10.1016/j.jss.2022.111359

Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of big data*, *2*, 1-21. https://doi.org/10.1186/s40537-014-0007-7

Nalbantis, I. (2008). Evaluation of a hydrological drought index. *European water*, *23*(24), 67-77.

Nalbantis, I., & Tsakiris, G. (2009). Assessment of hydrological drought revisited. *Water Resources Management*, *23*(5), 881–897. https://doi.org/10.1007/s11269-008-9305-1

Nandgude, N., Singh, T. P., Nandgude, S., & Tiwari, M. (2023). Drought prediction: A comprehensive review of different drought prediction models and adopted technologies. *Sustainability (Switzerland)*, *15*(15). https://doi.org/10.3390/su151511684

Naumann, G., Cammalleri, C., Mentaschi, L., & Feyen, L. (2021). Increased economic drought impacts in Europe with anthropogenic warming. *Nature Climate Change*, *11*(6), 485–491. https://doi.org/10.1038/s41558-021-01044-3

Ng, A., & Yap, N. T. (2011). Drought preparedness and response as if development matters: Case studies from Kenya. *Journal of Rural and Community Development*, *6*(2). https://journals.brandonu.ca/jrcd/article/view/512

Nguyen, N. A. (2017). Historic drought and salinity intrusion in the Mekong Delta in 2016: Lessons learned and response solutions. *Vietnam Journal of Science and Technology*, *1*(1), 2015-2018.

Nikulin, G., Jones, C., Giorgi, F., Asrar, G., Büchner, M., Cerezo-Mota, R., Christensen, O. B., Déqué, M., Fernandez, J., Hänsler, A., van Meijgaard, E., Samuelsson, P., Sylla, M. B., & Sushama, L. (2012). Precipitation climatology in an ensemble of CORDEX-Africa regional climate simulations. *Journal of Climate*, *25*(18), 6057-6078. https://doi.org/10.1175/JCLI-D-11-00375.1

Ning, Y., Liang, G., Ding, W., Shi, X., Fan, Y., Chang, J., Wang, Y., He, B., & Zhou, H. (2022). A mutual information theory - based approach for assessing uncertainties in deterministic multi - category precipitation forecasts. *Water Resources Research*, *58*(11), e2022WR032631. https://doi.org/10.1029/2022WR032631

Niu, Y., Wang, Y. Da, Mostaghimi, P., Swietojanski, P., & Armstrong, R. T. (2020). An innovative application of generative adversarial networks for physically accurate rock images with an unprecedented field of view. *Geophysical Research Letters*, *47*(23), 1–11. https://doi.org/10.1029/2020GL089029

Nogueira, A. R., Pugnana, A., Ruggieri, S., Pedreschi, D., & Gama, J. (2022). Methods and tools for causal discovery and causal inference. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *12*(2), 1–39. https://doi.org/10.1002/widm.1449

Novick, K. A., Biederman, J. A., Desai, A. R., Litvak, M. E., Moore, D. J. P., Scott, R. L., & Torn, M. S. (2018). The AmeriFlux network: A coalition of the willing. *Agricultural and Forest Meteorology*, *249*(May 2017), 444–456. https://doi.org/10.1016/j.agrformet.2017.10.009

Nowack, P., Runge, J., Eyring, V., Haigh, J.D. (2020). Causal networks for climate model evaluation and constrained projections. *Nature Communications*, *11*, 1415. https://doi.org/10.1038/s41467-020-15195-y.

Oikonomou, P. D., Karavitis, C. A., Tsesmelis, D. E., Kolokytha, E., & Maia, R. (2020). Drought characteristics assessment in Europe over the past 50 years. *Water Resources Management*, *34*(15), 4757–4772. https://doi.org/10.1007/s11269-020-02688-0

Okujeni, A., Kowalski, K., Lewińska, K. E., Schneidereit, S., & Hostert, P. (2024). Multidecadal grassland fractional cover time series retrieval for Germany from the Landsat and Sentinel-2 archives. *Remote Sensing of Environment*, *302*(November 2023). https://doi.org/10.1016/j.rse.2023.113980

Ombadi, M., Nguyen, P., Sorooshian, S., & Hsu, K. lin. (2020). Evaluation of methods for causal discovery in hydrometeorological systems. *Water Resources Research*, *56*(7), 1–22. https://doi.org/10.1029/2020WR027251

Ordóñez, P., Ribera, P., Gallego, D., & Peña-Ortiz, C. (2012). Major moisture sources for Western and Southern India and their role on synoptic-scale rainfall events. *Hydrological Processes*, *26*(25), 3886-3895. https://doi.org/10.1002/hyp.8455

Otkin, J. A., Woloszyn, M., Wang, H., Svoboda, M., Skumanich, M., Pulwarty, R., Lisonbee, J., Hoell, A., Hobbins, M., Haigh, T., & Cravens, A. E. (2022). Getting ahead of flash drought: From early warning to early action. *Bulletin of the American Meteorological Society*, *103*(10), E2188-E2202. https://doi.org/10.1175/BAMS-D-21-0288.1

Ouyang, H., Wei, X., & Wu, Q. (2019). Agricultural commodity futures prices prediction via long-and short-term time series network. *Journal of Applied Economics*, *22*(1), 468-483. https://doi.org/10.1080/15140326.2019.1668664

Owen, T., Svoboda, M., & Pulwarty, R. (2007,). National Integrated Drought Information System (NIDIS), United State Drought Portal (USDP): A window on drought information, impacts and implications. In *AGU Fall Meeting Abstracts* (Vol. 2007, pp. H23F-1684).

Özger, M., Mishra, A. K., & Singh, V. P. (2012). Long lead time drought forecasting using a wavelet and fuzzy logic combination model: a case study in Texas. *Journal of Hydrometeorology*, *13*(1), 284-297. https://doi.org/10.1175/JHM-D-10-05007.1

Pahlevan, N., Chittimalli, S. K., Balasubramanian, S. V., & Vellucci, V. (2019). Sentinel-2/Landsat-8 product consistency and implications for monitoring aquatic systems. *Remote sensing of Environment*, *220*, 19-29. https://doi.org/10.1016/j.rse.2018.10.027

Palmer, W. C. (1965). *Meteorological drought* (Vol. 30). US Department of Commerce, Weather Bureau.

Palmer, W. C. (1968). Keeping track of crop moisture conditions, nationwide: The new crop moisture index. *Weatherwise*, *21*, 156-161.

Pan, Y., Zhu, Y., Lü, H., Yagci, A. L., Fu, X., Liu, E., Xu, H., Ding, Z., & Liu, R. (2023). Accuracy of agricultural drought indices and analysis of agricultural drought characteristics in China between 2000 and 2019. *Agricultural Water Management*, *283*(April), 108305. https://doi.org/10.1016/j.agwat.2023.108305

Park, H., Kim, K., & Lee, D. K. (2019). Prediction of severe drought area based on random forest: Using satellite image and topography data. *Water (Switzerland)*, *11*(4), 1–15. https://doi.org/10.3390/w11040705

Park, S., Im, J., Jang, E., & Rhee, J. (2016). Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions. *Agricultural and Forest Meteorology*, *216*, 157–169. https://doi.org/10.1016/j.agrformet.2015.10.011

Pascanu, R., Gulcehre, C., Cho, K., & Bengio, Y. (2014). How to construct deep recurrent neural networks. *2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings*, 1–13.

Pathak, A., Ghosh, S., Alejandro Martinez, J., Dominguez, F., & Kumar, P. (2017). Role of oceanic and land moisture sources and transport in the seasonal and interannual variability of summer monsoon in India. *Journal of Climate*, *30*(5), 1839–1859. https://doi.org/10.1175/JCLI-D-16-0156.1

Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, *82*(4), 669-688. https://doi.org/10.1093/biomet/82.4.669

Pearl, J. (2000). Models, reasoning and inference. *Cambridge, UK: Cambridge University Press*, *19*(2), 3.

Pearl, J. (2010). Causal inference. *Causality: objectives and assessment*, 39-58.

Pearl, J., Glymour, M., & Jewell, N. P. (2016). *Causal inference in statistics: A primer*. John Wiley & Sons.

Peixóto, J. P., & Oort, A. H. (1984). Physics of climate. *Reviews of Modern Physics*, *56*(3), 365. https://doi.org/10.1103/RevModPhys.56.365

Peña-Gallardo, M., Vicente-Serrano, S. M., Hannaford, J., Lorenzo-Lacruz, J., Svoboda, M., Domínguez-Castro, F., Maneta, M., Tomas-Burguera, M., & El Kenawy, A. (2019). Complex influences of meteorological drought time-scales on hydrological droughts in natural basins of the contiguous Unites States. *Journal of Hydrology*, *568*, 611-625. https://doi.org/10.1016/j.jhydrol.2018.11.026

Phan, V. H., Dinh, V. T., & Su, Z. (2020). Trends in long-term drought changes in the Mekong River Delta of Vietnam. *Remote Sensing*, *12*(18), 1–22. https://doi.org/10.3390/RS12182974

Pinto, J. G., Ulbrich, S., Parodi, A., Rudari, R., Boni, G., & Ulbrich, U. (2013). Identification and ranking of extraordinary rainfall events over Northwest Italy: The role of Atlantic moisture. *Journal of Geophysical Research Atmospheres*, *118*(5), 2085–2097. https://doi.org/10.1002/jgrd.50179

Pokhrel, Y., Felfelani, F., Satoh, Y., Boulange, J., Burek, P., Gädeke, A., Gerten, D., Gosling, S. N., Grillakis, M., Gudmundsson, L., Hanasaki, N., Kim, H., Koutroulis, A., Liu, J., Papadimitriou, L., Schewe, J., Müller Schmied, H., Stacke, T., Telteu, C. E., Thiery, W., Veldkamp, T., Zhao, F., & Wada, Y. (2021). Global terrestrial water storage and drought severity under climate change. *Nature Climate Change*, *11*(3), 226–233. https://doi.org/10.1038/s41558-020-00972-w

Poornima, S., & Pushpalatha, M. (2019). Drought prediction based on SPI and SPEI with varying timescales using LSTM recurrent neural network. *Soft Computing*, *23*(18), 8399–8412. https://doi.org/10.1007/s00500-019-04120-1

Pozzi, W., Sheffield, J., Stefanski, R., Cripe, D., Pulwarty, R., Vogt, J. V., Heim, R. R., Jr., Brewer, M. J., Svoboda, M., Westerhoff, R., van Dijk, A. I. J. M., Lloyd-Hughes, B., Pappenberger, F., Werner, M., Dutra, E., Wetterhall, F., Wagner, W., Schubert, S., Mo, K., Nicholson, M., Bettio, L., Nunez, L., van Beek, R., Bierkens, M., de Goncalves, L. G. G., de Mattos, J. G. Z., & Lawford, R. (2013). Toward global drought early warning capability: Expanding international cooperation for the development of a framework for monitoring and forecasting. *Bulletin of the American Meteorological Society*, *94*(6), 776-785. https://doi.org/10.1175/BAMS-D-11-00176.1

Prudente, V. H. R., Martins, V. S., Vieira, D. C., e Silva, N. R. D. F., Adami, M., & Sanches, I. D. A. (2020). Limitations of cloud cover for optical remote sensing of agricultural areas across South America. *Remote Sensing Applications: Society and Environment*, *20*, 100414. https://doi.org/10.1016/j.rsase.2020.100414

Pulwarty, R. S., & Sivakumar, M. V. (2014). Information systems in a changing climate: Early warnings and drought risk management. *Weather and Climate Extremes*, *3*, 14-21. https://doi.org/10.1016/j.wace.2014.03.005

Racah, E., Beckham, C., Maharaj, T., Kahou, S. E., Prabhat, & Pal, C. (2017). ExtremeWeather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events. *Advances in Neural Information Processing Systems*, *2017-December*(Nips), 3403–3414.

Rahman, K. U., Shang, S., & Zohaib, M. (2021). Assessment of merged satellite precipitation datasets in monitoring meteorological drought over pakistan. *Remote Sensing*, *13*(9), 1–37. https://doi.org/10.3390/rs13091662

Ramsey, J., Zhang, K., Glymour, M., Romero, Sanchez Ruben Huang, B., Ebert-Uphoff, I., Samarasinghe, S., Barnes, E. A., & Clark, G. (2018). Tetrad - a toolbox for causal discovery. *8th International Workshop on Climate Informatics*, 4. https://www.engr.colostate.edu/~iebert/PAPERS/CI2018_paper_35.pdf

Rana, S., Mcgregor, J., & Renwick, J. (2015). Precipitation seasonality over the Indian subcontinent: An evaluation of gauge, reanalyses, and satellite retrievals. *Journal of Hydrometeorology*, *16*(2), 631–651. https://doi.org/10.1175/JHM-D-14-0106.1

Rasp, S., Hoyer, S., Merose, A., Langmore, I., Battaglia, P., Russell, T., Sanchez-Gonzalez, A., Yang, V., Carver, R., Agrawal, S., Chantry, M., Ben Bouallegue, Z., Dueben, P., Bromberg, C., Sisk, J., Barrington, L., Bell, A., & Sha, F. (2024). WeatherBench 2: A benchmark for the next generation of data-driven global weather models. *Journal of Advances in Modeling Earth Systems*, *16*(6). https://doi.org/10.1029/2023MS004019

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1

Reid, W. V., Chen, D., Goldfarb, L., Hackmann, H., Lee, Y. T., Mokhele, K., Ostrom, E., Raivio, K., Rockström, J., Schellnhuber, H. J., & Whyte, A. (2010). Earth system science for global sustainability: grand challenges. *Science*, *330*(6006), 916-917. https://doi.org/10.1126/science.1196263

Rhee, J., & Im, J. (2017). Meteorological drought forecasting for ungauged areas based on machine learning: Using long-range climate forecast and remote sensing data. *Agricultural and Forest Meteorology*, *237–238*, 105–122. https://doi.org/10.1016/j.agrformet.2017.02.011

Rienecker, M. M., Suarez, M. J., Todling, R., Bacmeister, J., Takacs, L., Liu, H. C., Gu, W., Sienkiewicz, M., Koster, R.D., Gelaro, R., Stajner, I., & Nielsen, J. E. (2008). *The GEOS-5 data assimilation system-documentation of versions 5.0. 1, 5.1. 0, and 5.2. 0* (No. NASA/TM-2008-104606-VOL-27).

Rios-Entenza, A., Soares, P. M., Trigo, R. M., Cardoso, R. M., & Miguez-Macho, G. (2014). Moisture recycling in the Iberian Peninsula from a regional climate simulation: Spatiotemporal analysis and impact on the precipitation regime. *Journal of Geophysical Research: Atmospheres*, *119*(10), 5895-5912. https://doi.org/10.1002/2013JD021274

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., & Toll, D. (2004). The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society*, *85*(3), 381–394. https://doi.org/10.1175/BAMS-85-3-381

Rojas, O., Vrieling, A., & Rembold, F. (2011). Assessing drought probability for agricultural areas in Africa with coarse resolution remote sensing imagery. *Remote Sensing of Environment*, *115*(2), 343–352. https://doi.org/10.1016/j.rse.2010.09.006

Rossi, G., & Cancelliere, A. (2013). Managing drought risk in water supply systems in Europe: A review. *International Journal of Water Resources Development*, *29*(2), 272–289. https://doi.org/10.1080/07900627.2012.713848

Rouholahnejad Freund, E., Zappa, M., & Kirchner, J. W. (2020). Averaging over spatiotemporal heterogeneity substantially biases evapotranspiration rates in a mechanistic large-scale land evaporation model. *Hydrology and Earth System Sciences*, *24*, 5015-5025. https://doi.org/10.5194/hess-24-5015-2020

Roundy, J. K., Ferguson, C. R., & Wood, E. F. (2014). Impact of land-atmospheric coupling in CFSv2 on drought prediction. *Climate Dynamics*, *43*(1–2), 421–434. https://doi.org/10.1007/s00382-013-1982-7

Roy, T., Martinez, J. A., Herrera-Estrada, J. E., Zhang, Y., Dominguez, F., Berg, A., Ek, M., & Wood, E. F. (2019). Role of moisture transport and recycling in characterizing droughts: Perspectives from two recent U.S. droughts and the CFSv2 system. *Journal of Hydrometeorology*, *20*(1), 139-154. https://doi.org/10.1175/JHM-D-18-0159.1

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688–701. https://doi.org/10.1037/h0037350

Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, *100*(469), 322-331. https://doi.org/10.1198/016214504000001880322

Ruehr, N. K., Grote, R., Mayr, S., & Arneth, A. (2019). Beyond the extreme: recovery of carbon and water relations in woody plants following heat and drought stress. *Tree physiology*, *39*(8), 1285-1299. https://doi.org/10.1093/treephys/tpz032

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533-536. https://doi.org/10.1038/323533a0

Runge, J. (2018). Causal network reconstruction from time series: from theoretical assumptions to practical estimation. *Chaos*, *28*(7), 075310. https://doi.org/10.1063/1.5025050.

Runge, J. (2020). Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets. *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence, UAI 2020*, *124*, 1388–1397.

Runge, J. (2022). Tigramite – Causal inference for time series datasets: Version 5.0 [Software]. *GitHub*. https://github.com/jakobrunge/tigramite

Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., Glymour, C., Kretschmer, M., Mahecha, M. D., Muñoz-Marí, J., van Nes, E. H., Peters, J., Quax, R., Reichstein, M., Scheffer, M., Schölkopf, B., Spirtes, P., Sugihara, G., Sun, J., Zhang, K., & Zscheischler, J. (2019a). Inferring causation from time series in Earth system sciences. *Nature Communications*, *10*(1), 1–13. https://doi.org/10.1038/s41467-019-10105-3

Runge, J., Gerhardus, A., Varando, G., Eyring, V., & Camps-Valls, G. (2023). Causal inference for time series. *Nature Reviews Earth & Environment*, *4*(7), 487-505. https://doi.org/10.1038/s43017-023-00431-y

Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., & Sejdinovic, D. (2019b). Detecting and quantifying causal associations in large nonlinear time series datasets. *Science Advances*, *5*(11). https://doi.org/10.1126/sciadv.aau4996

Saha, S., Moorthi, S., Pan, H., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y., Chuang, H., Juang, H. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., van den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., & Goldberg, M. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, *91*(8), 1015–1057. https://doi.org/10.1175/2010BAMS3001.1

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y., Chuang, H., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P., van den Dool, H., Zhang, Q., Wang, W., Chen, M., & Becker, E. (2014). The NCEP Climate Forecast System version 2.

Journal of Climate, *27*(6), 2185-2208. https://doi.org/10.1175/JCLI-D-12-00823.1

Sahoo, A. K., Sheffield, J., Pan, M., & Wood, E. F. (2015). Evaluation of the Tropical Rainfall Measuring Mission Multi-Satellite Precipitation Analysis (TMPA) for assessment of large-scale meteorological drought. *Remote Sensing of Environment*, *159*, 181–193. https://doi.org/10.1016/j.rse.2014.11.032

Salvucci, G. D., & Gentine, P. (2013). Emergent relation between surface vapor conductance and relative humidity profiles yields evaporation rates from weather data. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(16), 6287–6291. https://doi.org/10.1073/pnas.1215844110

Sam, T. T., Khoi, D. N., Thao, N. T. T., Nhi, P. T. T., Quan, N. T., Hoan, N. X., & Nguyen, V. T. (2019). Impact of climate change on meteorological, hydrological and agricultural droughts in the Lower Mekong River Basin: a case study of the Srepok Basin, Vietnam. *Water and Environment Journal*, *33*(4), 547-559. https://doi.org/10.1111/wej.12424

Sánchez, N., González-Zamora, Á., Piles, M., & Martínez-Fernández, J. (2016). A new Soil Moisture Agricultural Drought Index (SMADI) integrating MODIS and SMOS products: A case of study over the Iberian Peninsula. *Remote Sensing*, *8*(4). https://doi.org/10.3390/rs8040287

Santanello, J. A., Dirmeyer, P. A., Ferguson, C. R., Findell, K. L., Tawfik, A. B., Berg, A., Ek, M., Gentine, P., Guillod, B. P., Van Heerwaarden, C., Roundy, J., & Wulfmeyer, V. (2018). Land-atmosphere interactions the LoCo perspective. *Bulletin of the American Meteorological Society*, *99*(6), 1253–1272. https://doi.org/10.1175/BAMS-D-17-0001.1

Santanello, J. A., Peters-Lidard, C. D., & Kumar, S. V. (2011). Diagnosing the sensitivity of local land-atmosphere coupling via the soil moisture-boundary layer interaction. *Journal of Hydrometeorology*, *12*(5), 766–786. https://doi.org/10.1175/JHM-D-10-05014.1

Santanello, J. A., Peters-Lidard, C. D., Kennedy, A., & Kumar, S. V. (2013). Diagnosing the nature of land-atmosphere coupling: A case study of dry/wet extremes in the U.S. southern great plains. *Journal of Hydrometeorology*, *14*(1), 3-24. https://doi.org/10.1175/JHM-D-12-023.1

Santanello, J. A., Peters-Lidard, C. D., Kumar, S. V., Alonge, C., & Tao, W. K. (2009). A modeling and observational framework for diagnosing local land-atmosphere coupling on diurnal time scales. *Journal of Hydrometeorology*, *10*(3), 577–599.

https://doi.org/10.1175/2009JHM1066.1

Sarker, I. H. (2021). Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN computer science*, *2*(6), 420. https://doi.org/10.1007/s42979-021-00815-1

Satoh, Y., Yoshimura, K., Pokhrel, Y., Kim, H., Shiogama, H., Yokohata, T., Hanasaki, N., Wada, Y., Burek, P., Byers, E., Schmied, H. M., Gerten, D., Ostberg, S., Gosling, S. N., Boulange, J. E. S., & Oki, T. (2022). The timing of unprecedented hydrological drought under climate change. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-30729-2

Satori, M. A., & Antsaklis, P. J. (1991). A simple method to derive the bounds on the size to train multilayer neural networks. *IEEE Transactions on Neural Networks*, *2*(4), 467-471. https://doi.org/10.1109/72.88168

Savenije, H. H. (1996). The runoff coefficient as the key to moisture recycling. *Journal of Hydrology*, *176*(1-4), 219-225. https://doi.org/10.1016/0022-1694(95)02776-9

Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., & Monfardini, G. (2009). The graph neural network model. *IEEE transactions on neural networks*, *20*(1), 61-80. https://doi.org/10.1109/TNN.2008.2005605

Schepen, A., Wang, Q. J., & Robertson, D. E. (2014). Seasonal forecasts of Australian rainfall through calibration and bridging of coupled GCM outputs. *Monthly Weather Review*, *142*(5), 1758–1770. https://doi.org/10.1175/MWR-D-13-00248.1

Schmitt, T., Eisenberg, J., & Rao, R. R. (Eds.). (2007). *Improving disaster management: the role of IT in mitigation, preparedness, response, and recovery*. National Academies Press.

Schumacher, D. L., Keune, J., Dirmeyer, P., & Miralles, D. G. (2022). Drought self-propagation in drylands due to land–atmosphere feedbacks. *Nature Geoscience*, *15*(4), 262-268. https://doi.org/10.1038/s41561-022-00912-7

Schumacher, D. L., Keune, J., Van Heerwaarden, C. C., Vilà-Guerau de Arellano, J., Teuling, A. J., & Miralles, D. G. (2019). Amplification of mega-heatwaves through heat torrents fuelled by upwind drought. *Nature Geoscience*, *12*(9), 712-717. https://doi.org/10.1038/s41561-019-0431-6

Schwingshackl, C., Hirschi, M., & Seneviratne, S. I. (2018). A theoretical approach to assess soil moisture-climate coupling across CMIP5 and GLACE-CMIP5 experiments. *Earth System Dynamics*, *9*(4), 1217–1234. https://doi.org/10.5194/esd-9-1217-2018

Seager, R., & Hoerling, M. (2014). Atmosphere and ocean origins of North American droughts. *Journal of Climate*, *27*(12), 4581-4606. https://doi.org/10.1175/JCLI-D-13-00329.1

Seager, R., Hoerling, M., Schubert, S., Wang, H., Lyon, B., Kumar, A., Nakamura, J., & Henderson, N. (2015). Causes of the 2011-14 California drought. *Journal of Climate*, *28*(18), 6997–7024. https://doi.org/10.1175/JCLI-D-14-00860.1

Seager, R., Naik, N., & Vecchi, G. A. (2010). Thermodynamic and dynamic mechanisms for large-scale changes in the hydrological cycle in response to global warming. *Journal of Climate*, *23*(17), 4651–4668. https://doi.org/10.1175/2010JCLI3655.1

Sebesvari, Z., Le, H. T. T., Van Toan, P., Arnold, U., & Renaud, F. G. (2012). Agriculture and water quality in the Vietnamese Mekong Delta. In *The Mekong Delta system: Interdisciplinary analyses of a river delta*, 331-361. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-3962-8_13

Sene, K. (2024). Drought early warning systems. In *Hydrometeorology: Forecasting and applications* (pp. 371-398). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-58269-1_10

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., & Teuling, A. J. (2010). Investigating soil moisture-climate interactions in a changing climate: A review. *Earth-Science Reviews*, *99*(3-4), 125-161. https://doi.org/10.1016/j.earscirev.2010.02.004

Seneviratne, S. I., Lüthi, D., Litschi, M., & Schär, C. (2006). Land-atmosphere coupling and climate change in Europe. *Nature*, *443*(7108), 205–209. https://doi.org/10.1038/nature05095

Seneviratne, S. I., Wilhelm, M., Stanelle, T., Van Den Hurk, B., Hagemann, S., Berg, A., Cheruy, F., Higgins, M. E., Meier, A., Brovkin, V., Claussen, M., Ducharne, A., Dufresne, J. L., Findell, K. L., Ghattas, J., Lawrence, D. M., Malyshev, S., Rummukainen, M., & Smith, B. (2013). Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment. *Geophysical Research Letters*, *40*(19), 5212–5217. https://doi.org/10.1002/grl.50956

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh, M., Vicente-Serrano, S. M., Wehner, M., & Zhou, B. (2021). Weather and climate extreme events in a Changing climate. In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the*

Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1513–1766, https://doi.org/10.1017/9781009157896.013.

Seo, Y. W., & Ha, K. J. (2022). Changes in land-atmosphere coupling increase compound drought and heatwaves over northern East Asia. *Npj Climate and Atmospheric Science*, *5*(1). https://doi.org/10.1038/s41612-022-00325-8

Shafer, B. A., & Dezman, L. E. (1982). Development of a surface water supply index (SWSI) to assess the severity of drought conditions in snowpack runoff areas. *Proceedings of the Western Snow Conference Fiftieth Annual Meeting*, 164-175.

Shah, R. D., & Mishra, V. (2015). Development of an experimental near-real-time drought monitor for India. *Journal of Hydrometeorology*, *16*(1), 327-345. https://doi.org/10.1175/JHM-D-14-0041.1

Shahpari, G., Sadeghi, H., Ashena, M., & García-León, D. (2022). Drought effects on the Iranian economy: a computable general equilibrium approach. *Environment, Development and Sustainability*, *24*(3), 4110-4127. https://doi.org/10.1007/s10668-021-01607-6

Sharma, J., Granmo, OC., Goodwin, M., & Fidje, J. T. (2017). Deep convolutional neural networks for fire detection in images (pp. 183-193). Cham: Springer. https://doi.org/10.1007/978-3-319-65172-9_16

Sharma, S., & Mehra, R. (2020). Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology images—a comparative insight. *Journal of digital imaging*, *33*(3), 632-654. https://doi.org/10.1007/s10278-019-00307-y

Sheffield, J., & Wood, E.F. (2011). Drought: Past problems and future scenarios (1st ed.). Routledge. https://doi.org/10.4324/9781849775250

Sheffield, J., Andreadis, K. M., Wood, E. F., & Lettenmaier, D. P. (2009). Global and continental drought in the second half of the twentieth century: Severity-area-duration analysis and temporal variability of large-scale events. *Journal of Climate*, *22*(8), 1962–1981. https://doi.org/10.1175/2008JCLI2722.1

Sheffield, J., Wood, E. F., & Roderick, M. L. (2012). Little change in global drought over the past 60 years. *Nature*, *491*(7424), 435–438. https://doi.org/10.1038/nature11575

Shen, C. (2018). A transdisciplinary review of deep learning research and its relevance for water resources scientists. *Water Resources Research*, *54*(11), 8558–8593. https://doi.org/10.1029/2018WR022643

Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., & Woo, W. C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, *28*.

Shi, Y., Davis, K. J., Duffy, C. J., & Yu, X. (2013). Development of a coupled land surface hydrologic model and evaluation at a critical zone observatory. *Journal of Hydrometeorology*, *14*(5), 1401–1420. https://doi.org/10.1175/JHM-D-12-0145.1

Shimizu, S., Hoyer, P. O., Hyvärinen, A., Kerminen, A., & Jordan, M. (2006). A linear non-Gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, *7*(10). https://dl.acm.org/doi/abs/10.5555/1248547.1248619

Shinde, P. P., & Shah, S. (2018). A review of machine learning and deep learning applications. *Proceedings - 2018 4th International Conference on Computing, Communication Control and Automation, ICCUBEA 2018*, 1–6. https://doi.org/10.1109/ICCUBEA.2018.8697857

Shojaie, A., & Fox, E. B. (2022). Granger causality: A review and recent advances. *Annual Review of Statistics and Its Application*, *9*, 289–319. https://doi.org/10.1146/annurev-statistics-040120-010930

Shukla, S., & Wood, A. W. (2008). Use of a standardized runoff index for characterizing hydrologic drought. *Geophysical research letters*, *35*(2). https://doi.org/10.1029/2007GL032487

Shukla, S., Roberts, J., Hoell, A., Funk, C. C., Robertson, F., & Kirtman, B. (2019). Assessing North American Multimodel Ensemble (NMME) seasonal forecast skill to assist in the early warning of anomalous hydrometeorological events over East Africa. *Climate Dynamics*, *53*(12), 7411–7427. https://doi.org/10.1007/s00382-016-3296-z

Siegmund, J., Bliefernicht, J., Laux, P., & Kunstmann, H. (2015). Toward a seasonal precipitation prediction system for West Africa: Performance of CFSv2 and high-resolution dynamical downscaling. *Journal of Geophysical Research*, *120*(15), 7316–7339. https://doi.org/10.1002/2014JD022692

Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–14.

Siqueira, M., Katul, G., & Porporato, A. (2009). Soil moisture feedbacks on convection triggers: The role of soil-plant hydrodynamics. *Journal of Hydrometeorology*, *10*(1), 96–112. https://doi.org/10.1175/2008JHM1027.1

Slater, L. J., Villarini, G., & Bradley, A. A. (2017). Weighting of NMME temperature and precipitation forecasts across Europe. *Journal of Hydrology*, *552*, 646–659. https://doi.org/10.1016/j.jhydrol.2017.07.029

Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., Allan, R., Yin, X., Vose, R., Titchner, H., Kennedy, J., Spencer, L. J., Ashcroft, L., Brönnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T. A., Crouthamel, R., Domínguez-Castro, F., Freeman, J. E., Gergis, J., Hawkins, E., Jones, P. D., Jourdain, S., Kaplan, A., Kubota, H., Blancq, F. L., Lee, T.C., Lorrey, A., Luterbacher, J., Maugeri, M., Mock, C. J., Moore, G. W. K., Przybylak, R., Pudmenzky, C., Reason, C., Slonosky, V. C., Smith, C. A., Tinz, B., Trewin, B., Valente, M. A., Wang, X. L., Wilkinson, C., Wood, K., & Wyszyński, P. (2019). Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, *145*(724), 2876–2908. https://doi.org/10.1002/qj.3598

Smith, A.B., & Matthews, J. L. (2015). Quantifying uncertainty and variable sensitivity within the US billion-dollar weather and climate disaster cost estimates. *Natural Hazards*, *77*, 1829-1851. https://doi.org/10.1007/s11069-015-1678-x

Smith, M. D., Wilkins, K. D., Holdrege, M. C., Wilfahrt, P., Collins, S. L., Knapp, A. K., Sala, O. E., Dukes, J. S., Phillips, R. P., Yahdjian, L., Gherardi, L. A., Ohlert, T., Beier, C., Fraser, L. H., Jentsch, A., Loik, M. E., Maestre, F. T., Power, S. A., Yu, Q., Felton, A. J., Munson, S. M., Luo, Y., Abdoli, H., Abedi, M., Alados, C. L., Alberti, J., Alon, M., An, H., Anacker, B., Anderson, M., Auge, H., Bachle, S., Bahalkeh, K., Bahn, M., Batbaatar, A., Bauerle, T., Beard, K. H., Behn, K., Beil, I., Biancari, L., Blindow, I., Bondaruk, V. F., Borer, E. T., Bork, E. W., Bruschetti, C. M., Byrne, K. M., Cahill Jr., J. F., Calvo, D. A., Carbognani, M., Cardoni, A., Carlyle, C. N., Castillo-Garcia, M., Chang, S. X., Chieppa, J., Cianciaruso, M. V., Cohen, O., Cordeiro, A. L., Cusack, D. F., Dahlke, S., Daleo, P., D'Antonio, C. M. Dietterich, L. H., S. Doherty, T., Dubbert, M., Ebeling, A., Eisenhauer, N., Fischer, F. M., Forte, T'ai G. W., Gebauer, T., Gozalo, B., Greenville, A. C., Guidoni-Martins, K. G., Hannusch, H. J., Vatsø H. S., Hautier, Y., ... & Sun, W. (2024). Extreme drought impacts have been underestimated in grasslands and shrublands globally. *Proceedings of the National Academy of Sciences*, *121*(4), e2309881120. https://doi.org/10.1073/pnas.2309881120

Sodemann, H., & Stohl, A. (2013). Moisture origin and meridional transport in atmospheric rivers and their association with multiple cyclones. *Monthly Weather Review*, *141*(8), 2850-2868. https://doi.org/10.1175/MWR-D-12-00256.1

Sodemann, H., & Zubler, E. (2010). Seasonal and inter-annual variability of the moisture sources for alpine precipitation during 1995-2002. *International Journal of Climatology*, *30*(7), 947–961. https://doi.org/10.1002/joc.1932

Son, N. T., Chen, C. F., Chen, C. R., Chang, L. Y., & Minh, V. Q. (2012). Monitoring agricultural drought in the lower mekong basin using MODIS NDVI and land surface temperature data. *International Journal of Applied Earth Observation and Geoinformation*, *18*(1), 417–427. https://doi.org/10.1016/j.jag.2012.03.014

Sorí, R., Gimeno-Sotelo, L., Nieto, R., Liberato, M. L. R., Stojanovic, M., Pérez-Alarcón, A., Fernández-Alvarez, J. C., & Gimeno, L. (2023). Oceanic and terrestrial origin of precipitation over 50 major world river basins: Implications for the occurrence of drought. *Science of the Total Environment*, *859*(September 2022), 160288. https://doi.org/10.1016/j.scitotenv.2022.160288

Sorí, R., Nieto, R., Drumond, A., Vicente-Serrano, S. M., & Gimeno, L. (2017). The atmospheric branch of the hydrological cycle over the Indus, Ganges, and Brahmaputra river basins. *Hydrology and Earth System Sciences*, *21*(12), 6379–6399. https://doi.org/10.5194/hess-21-6379-2017

Sousa, P. M., Trigo, R. M., Aizpurua, P., Nieto, R., Gimeno, L., & Garcia-Herrera, R. (2011). Trends and extremes of drought indices throughout the 20th century in the Mediterranean. *Natural Hazards and Earth System Science*, *11*(1), 33–51. https://doi.org/10.5194/nhess-11-33-2011

Späth, F., Rajtschan, V., Weber, T. K. D., Morandage, S., Lange, D., Abbas, S. S., Behrendt, A., Ingwersen, J., Streck, T., & Wulfmeyer, V. (2023). The land-atmosphere feedback observatory: a new observational approach for characterizing land-atmosphere feedback. In *Geoscientific Instrumentation, Methods and Data Systems* (Vol. 12, Issue 1, pp. 25–44). https://doi.org/10.5194/gi-12-25-2023

Spinoni, J., Barbosa, P., De Jager, A., McCormick, N., Naumann, G., Vogt, J. V., Magni, D., Masante, D., & Mazzeschi, M. (2019). A new global database of meteorological drought events from 1951 to 2016. *Journal of Hydrology: Regional Studies*, *22*(January), 100593. https://doi.org/10.1016/j.ejrh.2019.100593

Spirtes, P., & Glymour, C. (1991). An algorithm for fast recovery of sparse causal graphs. *Social Science Computer Review*, *9*(1), 62–72. https://doi.org/10.1177/089443939100900106.

Srivastava, A. N., Nemani, R., & Steinhaeuser, K. (Eds.). (2017). *Large-scale machine learning in the earth sciences*. CRC Press.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, *15*(1), 1929-1958.

Srivastava, P. K. (2017). Satellite soil moisture: Review of theory and applications in water resources. *Water Resources Management*, *31*(10), 3161–3176. https://doi.org/10.1007/s11269-017-1722-6

Stanke, C., Kerac, M., Prudhomme, C., Medlock, J., & Murray, V. (2013). Health effects of drought: a systematic review of the evidence. *PLoS currents*, *5*. https://doi.org/10.1371%2Fcurrents.dis.7a2cee9e980f91ad7697b570bcc4b004

Stohl, A., & James, P. (2004). A Lagrangian analysis of the atmospheric branch of the global water cycle: Part 1: Method description, validation, and demonstration for the August 2002 flooding in central Europe. *Journal of Hydrometeorology*, *5*(4), 656–678. https://doi.org/10.1175/1525-7541(2004)005<0656:ALAOTA>2.0.CO;2

Stohl, A., & James, P. (2005). A Lagrangian analysis of the atmospheric branch of the global water cycle. Part II: Moisture transports between earth’s ocean basins and river catchments. *Journal of Hydrometeorology*, *6*(6), 961–984. https://doi.org/10.1175/JHM470.1

Stojanovic, M., Drumond, A., Nieto, R., & Gimeno, L. (2017). Moisture transport anomalies over the Danube River basin during two drought events: A Lagrangian analysis. *Atmosphere, 8*(10), 1–16. https://doi.org/10.3390/atmos8100193

Stojanovic, M., Drumond, A., Nieto, R., & Gimeno, L. (2018). Anomalies in moisture supply during the 2003 drought event in Europe: A Lagrangian analysis. *Water (Switzerland)*, *10*(4). https://doi.org/10.3390/w10040467

Stojanovic, M., Nieto, R., Liberato, M. L. R., Sorí, R., Vázquez, M., & Gimeno, L. (2021).

Tracking the origins of moisture over Vietnam: The role of moisture sources and atmospheric drivers on seasonal hydroclimatic conditions. *International Journal of Climatology*, *41*(13), 5843-5861. https://doi.org/10.1002/joc.7156

Strzepek, K., Yohe, G., Neumann, J., & Boehlert, B. (2010). Characterizing changes in drought risk for the United States from climate change. *Environmental Research Letters*, *5*(4), 044012. https://doi.org/10.1088/1748-9326/5/4/044012

Suárez-Cetrulo, A. L., Cervantes, A., & Quintana, D. (2019). Incremental market behavior classification in presence of recurring concepts. *Entropy*, *21*(1), 25. https://doi.org/10.3390/e21010025

Suárez-Cetrulo, A. L., Quintana, D., & Cervantes, A. (2023). A survey on machine learning for recurring concept drifting data streams. *Expert Systems with Applications*, *213*, 118934. https://doi.org/10.1016/j.eswa.2022.118934

Sugihara, G., May, R., Ye, H., Hsieh, C. H., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, *338*(6106), 496–500. https://doi.org/10.1126/science.1227079

Sun, B., & Wang, H. (2014). Moisture sources of semiarid grassland in China using the lagrangian particle model FLEXPART. *Journal of Climate*, *27*(6), 2457–2474. https://doi.org/10.1175/JCLI-D-13-00517.1

Sun, B., & Wang, H. (2015). Analysis of the major atmospheric moisture sources affecting three sub-regions of East China. *International Journal of Climatology*, *35*(9), 2243–2257. https://doi.org/10.1002/joc.4145

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K. L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, *56*(1), 79–107. https://doi.org/10.1002/2017RG000574

Sun, R., Yuan, H., Liu, X., & Jiang, X. (2016). Evaluation of the latest satellite-gauge precipitation products and their hydrologic applications over the Huaihe River basin. *Journal of Hydrology*, *536*, 302–319. https://doi.org/10.1016/j.jhydrol.2016.02.054

Sun, Z., Sandoval, L., Crystal-Ornelas, R., Mousavi, S. M., Wang, J., Lin, C., Cristea, N., Tong, D., Carande, W. H., Ma, X., Rao, Y., Bednar, J. A., Tan, A., Wang, J., Purushotham, S., Gill, T. E., Chastang, J., Howard, D., Holt, B., Gangodagamage, C., Zhao, P., Rivas, P., Chester, Z., Orduz, J., & John, A. (2022). A review of Earth artificial intelligence. *Computers* *and Geosciences*, *159*(January), 105034. https://doi.org/10.1016/j.cageo.2022.105034

Sur, C., Hur, J., Kim, K., Choi, W., & Choi, M. (2015). An evaluation of satellite-based drought indices on a regional scale. *International Journal of Remote Sensing*, *36*(22), 5593-5612. https://doi.org/10.1080/01431161.2015.1101653

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., Stooksbury, D., Miskus, D., & Stephens, S. (2002). The drought monitor. *Bulletin of the American Meteorological Society*, *83*(8), 1181-1190. https://doi.org/10.1175/1520-0477-83.8.1181

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).

Tabari, H., Nikbakht, J., & Hosseinzadeh Talaee, P. (2013). Hydrological drought assessment in Northwestern Iran based on Streamflow Drought Index (SDI). *Water Resources Management*, *27*(1), 137–151. https://doi.org/10.1007/s11269-012-0173-3

Tahmasebi, P., Kamrava, S., Bai, T., & Sahimi, M. (2020). Machine learning in geo-and environmental sciences: From small to large scale. *Advances in Water Resources*, *142*, 103619. https://doi.org/10.1016/j.advwatres.2020.103619

Takens, F. (1981). Detecting strange attractors in turbulence. In *Dynamical systems and turbulence*, *Warwick 1980*, (pp. 366–381). Berlin, Heidelberg: Springer. https://doi.org/10.1007/BFb0091924

Takeuchi, W., Darmawan, S., Shofiyati, R., Khiem, M. V., Oo, K. S., Pimple, U., & Heng, S. (2015). Near-real time meteorological drought monitoring and early warning system for croplands in asia. In *Asian Conference on Remote Sensing 2015: Fostering Resilient Growth in Asia* (Vol. 1, pp. 171-178).

Tao, H., Borth, H., Fraedrich, K., Su, B., & Zhu, X. (2014). Drought and wetness variability in the Tarim River Basin and connection to large-scale atmospheric circulation. *International Journal of Climatology*, *34*(8), 2678–2684. https://doi.org/10.1002/joc.3867

Tapiador, F. J., Turk, F. J., Petersen, W., Hou, A. Y., García-Ortega, E., Machado, L. A. T., Angelis, C. F., Salio, P., Kidd, C., Huffman, G. J., & de Castro, M. (2012). Global precipitation measurement: Methods, datasets and applications. *Atmospheric Research*, *104–105*, 70–97. https://doi.org/10.1016/j.atmosres.2011.10.021

Tavakol, A., McDonough, K. R., Rahmani, V., Hutchinson, S. L., & Hutchinson, J. M. S. (2021). The soil moisture data bank: The ground-based, model-based, and satellite-based soil moisture data. *Remote Sensing Applications: Society and Environment*, *24*(May), 100649. https://doi.org/10.1016/j.rsase.2021.100649

Tawfik, A. B., & Dirmeyer, P. A. (2014). A process-based framework for quantifying the atmospheric preconditioning of surface-triggered convection. *Geophysical Research Letters*, *41*(1), 173-178. https://doi.org/10.1002/2013GL057984

Thomas, B. F., Famiglietti, J. S., Landerer, F. W., Wiese, D. N., Molotch, N. P., & Argus, D. F. (2017). GRACE groundwater drought index: Evaluation of California Central Valley groundwater drought. *Remote Sensing of Environment*, *198*, 384–392. https://doi.org/10.1016/j.rse.2017.06.026

Tian, Y., Xu, Y. P., & Wang, G. (2018). Agricultural drought prediction using climate indices based on support vector regression in Xiangjiang River basin. *Science of the Total Environment*, *622–623*, 710–720. https://doi.org/10.1016/j.scitotenv.2017.12.025

Toms, B. A., Barnes, E. A., & Ebert-Uphoff, I. (2020). Physically interpretable neural networks for the geosciences: Applications to earth system variability. *Journal of Advances in Modeling Earth Systems*, *12*(9), 1–20. https://doi.org/10.1029/2019MS002002

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B. Ö., Floury, N., Brown, M., Traver, I. N., Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L'Abbate, M., Croci, R., Pietropaolo, A., Huchler, M., & Rostan, F. (2012). GMES Sentinel-1 mission. *Remote Sensing of Environment*, *120*, 9–24. https://doi.org/10.1016/j.rse.2011.05.028

Tran, D. D., van Halsema, G., Hellegers, P. J. G. J., Hoang, L. P., & Ludwig, F. (2019). Long-term sustainability of the Vietnamese Mekong Delta in question: An economic assessment of water management alternatives. *Agricultural Water Management*, *223*(July), 105703. https://doi.org/10.1016/j.agwat.2019.105703

Trenberth, K. E., Fasullo, J. T., & Mackaro, J. (2011). Atmospheric moisture transports from ocean to land and global energy flows in reanalyses. *Journal of Climate*, *24*(18), 4907–4924. https://doi.org/10.1175/2011JCLI4171.1

Tripathy, K. P., & Mishra, A. K. (2024). Deep learning in hydrology and water resources disciplines: concepts, methods, applications, and research directions. *Journal of Hydrology*, *628*(August 2023), 130458. https://doi.org/10.1016/j.jhydrol.2023.130458

Trnka, M., Hlavinka, P., Možný, M., Semerádová, D., Štěpánek, P., Balek, J., Bartošová, L., Zahradníček, P., Bláhová, M., Skalák, P., Farda, A., Hayes, M., Svoboda, M. D., Wagner, W., Eitzinger, J., Fischer, M., & Zalud, Z. (2020). Czech Drought Monitor System for monitoring and forecasting of agricultural drought and drought impacts. Papers in Natural Resources. 1392. https://digitalcommons.unl.edu/natrespapers/1392

Tsymbal, A. (2004). The problem of concept drift: definitions and related work. *Computer Science Department*, *Trinity College Dublin*, *106*(2), 58.

Tufaner, F., & Özbeyaz, A. (2020). Estimation and easy calculation of the Palmer Drought Severity Index from the meteorological data by using the advanced machine learning algorithms. *Environmental Monitoring and Assessment*, *192,* 1-14. https://doi.org/10.1007/s10661-020-08539-0

Tuinenburg, O. A., & Staal, A. (2020). Tracking the global flows of atmospheric moisture and associated uncertainties. *Hydrology and Earth System Sciences*, *24*(5), 2419–2435. https://doi.org/10.5194/hess-24-2419-2020

Tuttle, S. E. and G. D. Salvucci (2017), Confounding factors in determining causal soil moisture-precipitation feedback, *Water Resources Research*, *53,* 5531–5544, https://doi.org/10.1002/2016WR019869.

Ullrich, P. A., Xu, Z., Rhoades, A. M., Dettinger, M. D., Mount, J. F., Jones, A. D., & Vahmani, P. (2018). California's drought of the future: A midcentury recreation of the exceptional conditions of 2012–2017. *Earth's Future*, *6*(11), 1568–1587. https://doi.org/10.1029/2018EF001007

Umphlett, N., Timlin, M. S., Fuchs, B., Ryan, W., Doesken, N., Angel, J., Kellner, O., Hillaker, H. J., Knapp, M., Lin, X., Foster, S., Andresen, J., Pollyea, A., Spoden, G., Guinan, P., Akyüz, A., Rogers, J. C., Edwards, L. M., & Bergantino, T. (2012). National Integrated Drought Information System Central U.S. 2012 Drought Assessment. Publications of the National Drought Mitigation Center. 1. https://digitalcommons.unl.edu/ndmcpub/1

Unganai, L. S., & Kogan, F. N. (1998). Drought monitoring and corn yield estimation in Southern Africa from AVHRR data. *Remote sensing of environment*, *63*(3), 219-232. https://doi.org/10.1016/S0034-4257(97)00132-6

UNISDR (2006). Global survey of early warning systems: An assessment of capacities, gaps and opportunities toward building a comprehensive global early warning system for all natural hazards. *Platf Promot Early Warn UNISDR—PPEW UN*, 2006.

UNISDR (2009). Drought risk reduction framework and practices: Contributing to the implementation of the Hyogo Framework for Action.

United Nations Office for Disaster Risk Reduction (2015). Sendai framework for disaster risk reduction 2015-2030. In *Third United Nations World Conference on Disaster Risk Reduction*, Sendai, Japan. http://www.undrr.org/quick/11409

United Nations Resident Coordinator in Viet Nam (2016). Viet Nam: Drought and saltwater intrusion situation update No. 1.

United Nations Resident Coordinator in Viet Nam (2020). Vietnam 2020 drought and saltwater intrusion flash update No. 2.

Vachon, R. W., Welker, J. M., White, J. W. C., & Vaughn, B. H. (2010). Monthly precipitation isoscapes ($\delta^{18}O$) of the United States: Connections with surface temperatures, moisture source conditions, and air mass trajectories. *Journal of Geophysical Research Atmospheres*, *115*(21), 1–17. https://doi.org/10.1029/2010JD014105

Vadez, V., Grondin, A., Chenu, K., Henry, A., Laplaze, L., Millet, E. J., & Carminati, A. (2024). Crop traits and production under drought. *Nature Reviews Earth & Environment*, *5*(3), 211-225. https://doi.org/10.1038/s43017-023-00514-w

Valentini, R. (Ed.). (2003). *Fluxes of carbon, water and energy of European forests* (Vol. 163). Springer Science & Business Media.

van den Hurk, B., Best, M., Dirmeyer, P., Pitman, A., Polcher, J., & Santanello, J. (2011). Acceleration of land surface model development over a decade of GLASS. *Bulletin of the American Meteorological Society*, *92*(12), 1593-1600. http://www.jstor.org/stable/26218611

Van Der Ent, R. J. (2014). A new view on the hydrological cycle over continents, (Doctoral dissertation). Retrieved from TU Delft. (https://repository.tudelft.nl/islandora/object/uuid:0ab824ee-6956-4cc3-b530-3245ab4f32be). Delft, Netherlands: Delft University of Technology.

Van Der Ent, R. J., & Savenije, H. H. G. (2011). Length and time scales of atmospheric moisture recycling. *Atmospheric Chemistry and Physics*, *11*(5), 1853–1863. https://doi.org/10.5194/acp-11-1853-2011

Van Der Ent, R. J., & Savenije, H. H. G. (2013). Oceanic sources of continental precipitation and the correlation with sea surface temperature. *Water Resources Research*, *49*(7), 3993–4004. https://doi.org/10.1002/wrcr.20296

Van Der Ent, R. J., Savenije, H. H. G., Schaefli, B., & Steele-Dunne, S. C. (2010). Origin and fate of atmospheric moisture over continents. *Water Resources Research*, *46*(9), 1–12. https://doi.org/10.1029/2010WR009127

Van Der Ent, R. J., Tuinenburg, O. A., Knoche, H. R., Kunstmann, H., & Savenije, H. H. G. (2013). Should we use a simple or complex model for moisture recycling and atmospheric moisture tracking? *Hydrology and Earth System Sciences*, *17*(12), 4869–4884. https://doi.org/10.5194/hess-17-4869-2013

Van Der Ent, R. J., Wang-Erlandsson, L., Keys, P. W., & Savenije, H. H. G. (2014). Contrasting roles of interception and transpiration in the hydrological cycle – Part 2: Moisture recycling. *Earth System Dynamics*, *5*(2), 471–489. https://doi.org/10.5194/esd-5-471-2014

Van der Molen, M. K., Dolman, A. J., Ciais, P., Eglin, T., Gobron, N., Law, B. E., Meir, P., Peters, W., Phillips, O. L., Reichstein, M., Chen, T., Dekker, S. C., Doubková, M., Friedl, M. A., Jung, M., van den Hurk, B. J. J. M., de Jeu, R. A. M., Kruijt, B., Ohta, T., Rebel, K. T., Plummer, S., Seneviratne, S.I., Sitch, S., Teuling, A. J., van der Werf, Wang, G. R., & G. (2011). Drought and ecosystem carbon cycling. *Agricultural and Forest Meteorology*, *151*(7), 765–773. https://doi.org/10.1016/j.agrformet.2011.01.018

van der Schrier, G., Barichivich, J., Briffa, K. R., & Jones, P. D. (2013). A scPDSI-based global data set of dry and wet spells for 1901–2009. *Journal of Geophysical Research: Atmospheres*, *118*(10), 4025-4048. https://doi.org/10.1002/jgrd.50355

van Dijk, A. I. J. M., Beck, H. E., Crosbie, R. S., De Jeu, R. A. M., Liu, Y. Y., Podger, G. M., Timbal, B., & Viney, N. R. (2013). The Millennium Drought in southeast Australia (2001-2009): Natural and human causes and implications for water resources, ecosystems, economy, and society. *Water Resources Research*, *49*(2), 1040–1057. https://doi.org/10.1002/wrcr.20123

Van Lanen, H. A., & Peters, E. (2000). Definition, effects and assessment of groundwater droughts. In *Drought and drought mitigation in Europe* (pp. 49-61). Dordrecht: Springer Netherlands.

Van Loon, A. F. (2015). Hydrological drought explained. *Wiley Interdisciplinary Reviews: Water*, *2*(4), 359–392. https://doi.org/10.1002/WAT2.1085

Van Loon, A. F., & Van Lanen, H. A. J. (2012). A process-based typology of hydrological drought. *Hydrology and Earth System Sciences*, *16*(7), 1915–1946. https://doi.org/10.5194/hess-16-1915-2012

van Rooy, M. P. (1965). A rainfall anomaly index independent of time and space. *Notos*, *14*, 43-48.

Van Zyl, K. (2006). A study on a disaster risk management plan for the South African agricultural sector. *Pretoria: Business Print Centre*.

Vandal, T., Kodra, E., Ganguly, S., Michaelis, A., Nemani, R., & Ganguly, A. R. (2018). Generating high resolution climate change projections through single image super-resolution: An abridged version. *IJCAI International Joint Conference on Artificial Intelligence*, *2018-July*, 5389–5393. https://doi.org/10.24963/ijcai.2018/759

Varian, H. R. (2016). Causal inference in economics and marketing. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(27), 7310–7315. https://doi.org/10.1073/pnas.1510479113

Vicca, S., Balzarolo, M., Filella, I., Granier, A., Herbst, M., Knohl, A., Longdoz, B., Mund, M., Nagy, Z., Pintér, K., Rambal, S., Verbesselt, J., Verger, A., Zeileis, A., Zhang, C., & Peñuelas, J. (2016). Remotely-sensed detection of effects of extreme droughts on gross primary production. *Scientific Reports*, *6*(June), 1–13. https://doi.org/10.1038/srep28269

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of climate*, *23*(7), 1696-1718. https://doi.org/10.1175/2009JCLI2909.1

Vicente‐Serrano, S. M., López‐Moreno, J. I., Gimeno, L., Nieto, R., Morán‐Tejeda, E., Lorenzo‐Lacruz, J., Beguería, S., & Azorin‐Molina, C. (2011). A multiscalar global evaluation of the impact of ENSO on droughts. *Journal of Geophysical Research: Atmospheres*, *116*(D20). https://doi.org/10.1029/2011JD016039

Vogel, M. M., Orth, R., Cheruy, F., Hagemann, S., Lorenz, R., van den Hurk, B. J. J. M., & Seneviratne, S. I. (2017). Regional amplification of projected changes in extreme temperatures strongly controlled by soil moisture-temperature feedbacks. *Geophysical*

Research Letters, *44*(3), 1511–1519. https://doi.org/10.1002/2016GL071235

Vu, D. T., Yamada, T., & Ishidaira, H. (2018). Assessing the impact of sea level rise due to climate change on seawater intrusion in Mekong Delta, Vietnam. *Water Science and Technology*, *77*(6), 1632-1639. https://doi.org/10.2166/wst.2018.038

Wakefield, R. A., Turner, D. D., & Basara, J. B. (2021). Evaluation of a land–atmosphere coupling metric computed from a ground-based infrared interferometer. *Journal of Hydrometeorology*, *22*(8), 2073-2087. https://doi.org/10.1175/JHM-D-20-0303.1

Wang, H. J., Sun, J. Q., Chen, H. P., Zhu, Y. L., Zhang, Y., Jiang, D. B., Lang, X. M., Fan, K., Yu, E. T., & Yang, S. (2012). Extreme climate in China: Facts, simulation and projection. *Meteorologische Zeitschrift*, *21*(3), 279–304. https://doi.org/10.1127/0941-2948/2012/0330

Wang, H., & Chen, H. (2012). Climate control for southeastern China moisture and precipitation: Indian or East Asian monsoon? *Journal of Geophysical Research Atmospheres*, *117*(12), 1–9. https://doi.org/10.1029/2012JD017734

Wang, J., Tang, Q., Chen, A., Tang, Y., Xu, X., Yun, X., Mu, M., Wright, N., & Chen, D. (2022). Impacts of Summer Monsoons on flood characteristics in the Lancang-Mekong River Basin. *Journal of Hydrology*, *604*(July 2021), 127256. https://doi.org/10.1016/j.jhydrol.2021.127256

Wang, T., Tu, X., Singh, V. P., Chen, X., Lin, K., & Zhou, Z. (2023). Drought prediction: Insights from the fusion of LSTM and multi-source factors. *Science of the Total Environment*, *902*(July), 166361. https://doi.org/10.1016/j.scitotenv.2023.166361

Wang, T., Tu, X., Singh, V. P., Chen, X., Lin, K., Lai, R., & Zhou, Z. (2022). Socioeconomic drought analysis by standardized water supply and demand index under changing environment. *Journal of Cleaner Production*, *347*(October 2021), 131248. https://doi.org/10.1016/j.jclepro.2022.131248

Wang, Y., & Yuan, X. (2022). Land-atmosphere coupling speeds up flash drought onset. *Science of the Total Environment*, *851*(August), 158109. https://doi.org/10.1016/j.scitotenv.2022.158109

Wang, Y., Yang, J., Chen, Y., De Maeyer, P., Li, Z., & Duan, W. (2018). Detecting the causal effect of soil moisture on precipitation using convergent cross mapping. *Scientific Reports*, *8*(1), 1–8. https://doi.org/10.1038/s41598-018-30669-2

Wei, J., & Dirmeyer, P. A. (2019). Sensitivity of land precipitation to surface evapotranspiration: a nonlocal perspective based on water vapor transport. *Geophysical Research Letters*, *46*(21), 12588–12597. https://doi.org/10.1029/2019GL085613

Wei, J., Dirmeyer, P. A., Bosilovich, M. G., & Wu, R. (2012). Water vapor sources for Yangtze River Valley rainfall: Climatology, variability, and implications for rainfall forecasting. *Journal of Geophysical Research Atmospheres*, *117*(5), 1–11. https://doi.org/10.1029/2011JD016902

Wells, N., Goddard, S., & Hayes, M. J. (2004). A self-calibrating Palmer drought severity index. *Journal of climate*, *17*(12), 2335-2351. https://doi.org/10.1175/1520-0442(2004)017%3C2335:ASPDSI%3E2.0.CO;2

Western Governors' Association (2004). Creating a drought early warning system for the 21st century: The National Integrated Drought Information System. https://www.drought.gov/sites/default/files/2020-06/200406_WGA_NIDIS_Report.pdf#page=3.05

Whang, S. E., & Lee, J. G. (2020). Data collection and quality challenges for deep learning. *Proceedings of the VLDB Endowment*, *13*(12), 3429-3432. https://doi.org/10.14778/3415478.3415562

Whang, S. E., Roh, Y., Song, H., & Lee, J. G. (2023). Data collection and quality challenges in deep learning: A data-centric ai perspective. *The VLDB Journal*, *32*(4), 791-813. https://doi.org/10.1007/s00778-022-00775-9

Wilhite, D. A., & Glantz, M. H. (1985). Understanding: The drought phenomenon: The role of definitions. *Water International*, *10*(3), 111–120. https://doi.org/10.1080/02508068508686328

WMO & GWP (2016). Handbook of Drought Indicators and Indices (M. Svoboda and B. A. Fuchs). Integrated Drought Management Programme (IDMP), Integrated Drought Management Tools and Guidelines Series 2. Geneva.

Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 3-19).

World Meteorological Organization (1992). International meteorological vocabulary, 2nd edition. Publication no. 182. World Meteorological Organization (WMO), Geneva.

Wu, H., & Wilhite, D. A. (2004). An operational agricultural drought risk assessment model for Nebraska, USA. *Natural Hazards*, *33*, 1-21.

https://doi.org/10.1023/B:NHAZ.0000034994.44357.75

Wu, H., Hayes, M. J., Wilhite, D. A., & Svoboda, M. D. (2005). The effect of the length of record on the standardized precipitation index calculation. *International Journal of Climatology*, *25*(4), 505–520. https://doi.org/10.1002/joc.1142

Wu, H., Su, X., Singh, V. P., & Zhang, T. (2022). Predicting hydrological drought with Bayesian Model Averaging Ensemble Vine Copula (BMAViC) Model. *Water Resources Research*, *58*(11). https://doi.org/10.1029/2022WR033146

Wu, H., Su, X., Singh, V. P., AghaKouchak, A., & Liu, Z. (2023). Bayesian vine copulas improve agricultural drought prediction for long lead times. *Agricultural and Forest Meteorology*, *331*(May 2022), 109326. https://doi.org/10.1016/j.agrformet.2023.109326

Wu, H., Su, X., Singh, V. P., Feng, K., & Niu, J. (2021). Agricultural drought prediction based on conditional distributions of vine copulas. *Water Resources Research*, *57*(8), 1–23. https://doi.org/10.1029/2021WR029562

Wu, J., Wang, G., Chen, X., Yuan, X., Yao, H., Chen, X., Lan, T., Zheng, Y., & AghaKouchak, A. (2024). Hydrological drought characterization considering onset, maximum streamflow deficit, and termination. *Advances in Water Resources*, *184*(December 2023), 104613. https://doi.org/10.1016/j.advwatres.2023.104613

Wu, J., Zhou, Y., Wang, H., Wang, X., & Wang, J. (2023). Assessing the causal effects of climate change on vegetation dynamics in northeast China using convergence cross-mapping. *IEEE Access*, *11*(September), 115367–115379. https://doi.org/10.1109/ACCESS.2023.3325485

Wu, Y., Miao, C., Slater, L., Fan, X., Chai, Y., & Sorooshian, S. (2024). Hydrological projections under CMIP5 and CMIP6: Sources and magnitudes of uncertainty. *Bulletin of the American Meteorological Society*, *105*(1), E59-E74.

Wu, Z., Yin, H., He, H., & Li, Y. (2022). Dynamic-LSTM hybrid models to improve seasonal drought predictions over China. *Journal of Hydrology*, *615*(PA), 128706. https://doi.org/10.1016/j.jhydrol.2022.128706

Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R., & Woodcock, C. E. (2012). Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, *122*, 2–10. https://doi.org/10.1016/j.rse.2012.01.010

Wulfmeyer, V., Turner, D. D., Baker, B., Banta, R., Behrendt, A., Bonin, T., Brewer, W. A., Buban, M., Choukulkar, A., Dumas, E., Hardesty, R. M., Heus, T., Ingwersen, J., Lange, D., Lee, T. R., Metzendorf, S., Muppa, S. K., Meyers, T., Newsom, R., Osman, M., Raasch, S., Santanello, J., Senff, C., Späth, F., Wagner, T., & Weckwerth, T. (2018). A new research approach for observing and characterizing land atmosphere feedback. *Bulletin of the American Meteorological Society*, *99*(8), 1639–1667. https://doi.org/10.1175/BAMS-D-17-0009.1

Xiao, M., & Cui, Y. (2021). Source of evaporation for the seasonal precipitation in the Pearl River Delta, China. *Water Resources Research*, *57*(8), 1–12. https://doi.org/10.1029/2020WR028564

xmingzh. (2021). xmingzh/Modified-WAM2layers: First release (v1.0.0) [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.4796962

Xu, L., Chen, N., & Zhang, X. (2018b). A comparison of large-scale climate signals and the North American Multi-Model Ensemble (NMME) for drought prediction in China. *Journal of Hydrology*, *557*, 378–390. https://doi.org/10.1016/j.jhydrol.2017.12.044

Xu, L., Chen, N., Zhang, X., & Chen, Z. (2018a). An evaluation of statistical, NMME and hybrid models for drought prediction in China. *Journal of Hydrology*, *566*(January), 235–249. https://doi.org/10.1016/j.jhydrol.2018.09.020

Xu, L., Chen, N., Zhang, X., Chen, Z., Hu, C., & Wang, C. (2019). Improving the North American multi-model ensemble (NMME) precipitation forecasts at local areas using wavelet and machine learning. *Climate Dynamics*, *53*(1–2), 601–615. https://doi.org/10.1007/s00382-018-04605-z

Xu, J., Hu, C., Yang, S., Lian, T., & Cai, W. (2025). Unfolding the differences in two types of ENSO indices with various definitions: A systematic review with the same criteria. *Earth-Science Reviews*, 105076. https://doi.org/10.1016/j.earscirev.2025.105076

Yan, H., Moradkhani, H., & Zarekarizi, M. (2017). A probabilistic drought forecasting framework: A combined dynamical and statistical approach. *Journal of Hydrology*, *548*, 291–304. https://doi.org/10.1016/j.jhydrol.2017.03.004

Yanai, M., & Tomita, T. (1998). Seasonal and interannual variability of atmospheric heat sources and moisture sinks as determined from NCEP–NCAR reanalysis. *Journal of Climate*, *11*(3), 463-482. https://doi.org/10.1175/1520-

0442(1998)011%3C0463:SAIVOA%3E2.0.CO;2

Yang, Z., Varble, A., Berg, L. K., Qian, Y., Tai, S.-L., Chen, J., Gustafson Jr, W. I., Feng, Z., & Liu, Y. (2023). Sensitivity of precipitation displacement of a simulated MCS to changes in land Surface conditions. *Journal of Geophysical Research: Atmospheres*, 128, e2022JD037642. https://doi.org/10.1029/2022JD037642

Yao, L., Chu, Z., Li, S., Li, Y., Gao, J., & Zhang, A. (2021). A survey on causal inference. *ACM Transactions on Knowledge Discovery from Data*, *15*(5). https://doi.org/10.1145/3444944

Yeh, S. W., Cai, W., Min, S. K., McPhaden, M. J., Dommenget, D., Dewitte, B., Collins, M., Ashok, K., An, S., Yim, Y. B., & Kug, J. S. (2018). ENSO atmospheric teleconnections and their response to greenhouse gas forcing. *Reviews of Geophysics*, *56*(1), 185-206. https://doi.org/10.1002/2017RG000568

Yihdego, Y., Vaheddoost, B., & Al-Weshah, R. A. (2019). Drought indices and indicators revisited. *Arabian Journal of Geosciences*, *12*(3). https://doi.org/10.1007/s12517-019-4237-z

Yoon, J. H., Ruby Leung, L., & Correia Jr, J. (2012). Comparison of dynamically and statistically downscaled seasonal climate forecasts for the cold season over the United States. *Journal of Geophysical Research: Atmospheres*, *117*(D21). https://doi.org/10.1029/2012JD017650

Yu, S., & Ma, J. (2021). Deep learning for geophysics: Current and future trends. *Reviews of Geophysics*, *59*(3), 1–36. https://doi.org/10.1029/2021RG000742

Yuan, G., Zhang, L., & Liu, Y. (2021). Impacts of soil moisture and atmospheric moisture transport on the precipitation in two typical regions of China. *Atmospheric Research*, *247*(February 2020), 105151. https://doi.org/10.1016/j.atmosres.2020.105151

Yuan, X., Ma, Z., Pan, M., & Shi, C. (2015). Microwave remote sensing of short-term droughts during crop growing seasons. *Geophysical Research Letters*, *42*(11), 4394–4401. https://doi.org/10.1002/2015GL064125

Yukimoto, S., Koshiro, T., Kawai, H., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., Yoshimura, H., Shindo, E., Mizuta, R., Ishii, M., Obata, A., Adachi, Y. (2019). MRI MRI-ESM2.0 model output prepared for CMIP6 ScenarioMIP ssp585 [Dataset]. *Earth System Grid Federation*. https://doi.org/10.22033/ESGF/CMIP6.6929

Zargar, A., Sadiq, R., Naser, B., & Khan, F. I. (2011). A review of drought indices. *Environmental Reviews*, *19*(1), 333–349. https://doi.org/10.1139/a11-013

Zaytar, M. A., & El Amrani, C. (2016). Sequence to sequence weather forecasting with long short-term memory recurrent neural networks. *International Journal of Computer Applications*, *143*(11), 7-11.

Zeng, D., & Yuan, X. (2021). Modeling the influence of upstream land-atmosphere coupling on the 2017 persistent drought over northeast china. *Journal of Climate*, *34*(12), 4971–4988. https://doi.org/10.1175/JCLI-D-20-0650.1

Zeng, D., Yuan, X., & Roundy, J. K. (2019). Effect of teleconnected land-atmosphere coupling on northeast China persistent drought in spring-summer of 2017. *Journal of Climate*, *32*(21), 7403–7420. https://doi.org/10.1175/JCLI-D-19-0175.1

Zeng, J., Zhou, T., Qu, Y., Bento, V. A., Qi, J., Xu, Y., Li, Y., & Wang, Q. (2023). An improved global vegetation health index dataset in detecting vegetation drought. *Scientific Data*, *10*(1), 1–12. https://doi.org/10.1038/s41597-023-02255-3

Zhang, C. (2020). Moisture sources for precipitation in Southwest China in summer and the changes during the extreme droughts of 2006 and 2011. *Journal of Hydrology*, *591*(July), 125333. https://doi.org/10.1016/j.jhydrol.2020.125333

Zhang, C., Brodeur, Z. P., Steinschneider, S., & Herman, J. D. (2022). Leveraging spatial patterns in precipitation forecasts using deep learning to support regional water management. *Water Resources Research*, *58*(9), 1–18. https://doi.org/10.1029/2021WR031910

Zhang, C., Tang, Q., Chen, D., Li, L., Liu, X., & Cui, H. (2017). Tracing changes in atmospheric moisture supply to the drying Southwest China. *Atmospheric Chemistry and Physics*, *17*(17), 10383–10393. https://doi.org/10.5194/acp-17-10383-2017

Zhang, C., Vinyals, O., Munos, R., & Bengio, S. (2018). A study on overfitting in deep reinforcement learning. *arXiv preprint arXiv:1804.06893*. https://doi.org/10.48550/arXiv.1804.06893

Zhang, L., Jiao, W., Zhang, H., Huang, C., & Tong, Q. (2017). Studying drought phenomena in the Continental United States in 2011 and 2012 using various drought indices. *Remote Sensing of Environment*, *190*, 96–106. https://doi.org/10.1016/j.rse.2016.12.010

Zhang, R., Chen, Z. Y., Xu, L. J., & Ou, C. Q. (2019). Meteorological drought forecasting based on a statistical model with machine learning techniques in Shaanxi province, China.

Science of the Total Environment, *665*, 338–346. https://doi.org/10.1016/j.scitotenv.2019.01.431

Zhang, Y., Hao, Z., Feng, S., Zhang, X., Xu, Y., & Hao, F. (2021). Agricultural drought prediction in China based on drought propagation and large-scale drivers. *Agricultural Water Management*, *255*(June), 107028. https://doi.org/10.1016/j.agwat.2021.107028

Zhang, Y., Long, M., Chen, K., Xing, L., Jin, R., Jordan, M. I., & Wang, J. (2023). Skilful nowcasting of extreme precipitation with NowcastNet. *Nature*, *619*(7970), 526–532. https://doi.org/10.1038/s41586-023-06184-4

Zhang, Y., Qiao, S., Ji, S., & Li, Y. (2020). DeepSite: bidirectional LSTM and CNN models for predicting DNA–protein binding. *International Journal of Machine Learning and Cybernetics*, *11*(4), 841–851. https://doi.org/10.1007/s13042-019-00990-x

Zhao, M., & Running, S. W. (2010). Drought-induced reduction in global. *Science*, *329*(5994), 940–943. https://doi.org/10.1126/science.1192666

Zhao, S., Cong, D., He, K., Yang, H., & Qin, Z. (2017). Spatial-temporal variation of drought in China from 1982 to 2010 based on a modified Temperature Vegetation Drought Index (mTVDI). *Scientific Reports*, *7*(1), 1–12. https://doi.org/10.1038/s41598-017-17810-3

Zheng, Y., Huang, B., Chen, W., Ramsey, J., Gong, M., Cai, R., Shimizu, S., Spirtes, P., & Zhang, K. (2024). Causal-learn: Causal discovery in python. *Journal of Machine Learning Research*, *25*(60), 1-8. https://www.jmlr.org/papers/v25/23-0970.html

Zhi, W., Feng, D., Tsai, W. P., Sterle, G., Harpold, A., Shen, C., & Li, L. (2021). From hydrometeorology to river water quality: Can a deep learning model predict dissolved oxygen at the continental scale? *Environmental Science and Technology*, *55*(4), 2357–2368. https://doi.org/10.1021/acs.est.0c06783

Zhou, K., & Shi, X. (2024). Understanding precipitation moisture sources and their dominant factors during droughts in the Vietnamese Mekong Delta. *Water Resources Research*, *60*, e2023WR035920. https://doi.org/10.1029/2023WR035920

Zhou, K., Li, J., Zhang, T., & Kang, A. (2021). The use of combined soil moisture data to characterize agricultural drought conditions and the relationship among different drought types in China. *Agricultural Water Management*, *243*, 106479. https://doi.org/10.1016/j.agwat.2020.106479

Zhou, K., Zheng, Y., Li, B., Dong, W., & Zhang, X. (2019). Forecasting different types of convective weather: A deep learning approach. *Journal of Meteorological Research*, *33*(5), 797–809. https://doi.org/10.1007/s13351-019-8162-6

Zhou, S., Park Williams, A., Berg, A. M., Cook, B. I., Zhang, Y., Hagemann, S., Lorenz, R., Seneviratne, S. I., & Gentine, P. (2019). Land–atmosphere feedbacks exacerbate concurrent soil drought and atmospheric aridity. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(38), 18848–18853. https://doi.org/10.1073/pnas.1904955116

Zhou, S., Williams, A. P., Lintner, B. R., Berg, A. M., Zhang, Y., Keenan, T. F., Cook, B. I., Hagemann, S., Seneviratne, S. I., & Gentine, P. (2021). Soil moisture–atmosphere feedbacks mitigate declining water availability in drylands. *Nature Climate Change*, *11*(1), 38–44. https://doi.org/10.1038/s41558-020-00945-z

Zhuang, Y., Fu, R., & Wang, H. (2020). Large-scale atmospheric circulation patterns associated with U.S. Great Plains warm season droughts revealed by self-organizing maps. *Journal of Geophysical Research: Atmospheres, 125*(5), 1–18. https://doi.org/10.1029/2019JD031460

Zscheischler, J., Michalak, A. M., Schwalm, C., Mahecha, M. D., Huntzinger, D. N., Reichstein, M., Berthier, G., Ciais, P., Cook, R. B., El-Masri, B., Huang, M., Ito, A., Jain, A., King, A., Lei, H., Lu, C., Mao, J., Peng, S., Poulter, B., Ricciuto, D., Shi, X., Tao, B., Tian, H., Viovy, N., Wang, W., Wei, Y., Yang, J., & Zeng, N. (2014). Impact of large-scale climate extremes on biospheric carbon fluxes: An intercomparison based on MsTMIP data. *Global Biogeochemical Cycles*, *28*(6), 585–600. https://doi.org/10.1002/2014GB004826