Advancing drought impact early warning systems for vegetation: root-zone soil moisture simulations and ensemble forecasts for southeastern Australia

A thesis submitted for the degree of Doctor of Philosophy of The Australian National University

Yizhi Li

July 2025

Fenner School of Environment and Society

Australian National University, Canberra, ACT, Australia

Declaration

This thesis is submitted in Standard Format, comprising chapters adapted from published journal articles. I declare that the research presented herein is my original work conducted during my candidature at The Australian National University, except where explicitly stated in the Statement of Contribution for multi-author papers included in the thesis.

Publication declaration

Chapter 4 has been published. Chapters 3, 5 and 6 are presented in manuscript form. Details of the published paper, including my contribution to the article, are provided below.

Chapter 4: **Li, Y.**, van Dijk, A.I.J.M., Tian, S., Renzullo, L.J. (2023). Skill and lead time of vegetation drought impact forecasts based on soil moisture observations. *Journal of Hydrology,* 620, 129420. https://doi.org/10.1016/j.jhydrol.2023.129420

Author	Nature of contribution	Extent of contribution in research	Extent of contribution in writing
Yizhi Li	Led the project design, conducted data analysis, drafted the manuscript, and handled all revisions.	80%	70%
Albert I.J.M. van Dijk	Contributed to project design and provided supervisory guidance and critical manuscript revisions.	10%	20%
Siyuan Tian	Assisted in project design and provided critical feedback on manuscript drafts.	5%	5%
Renzullo J. Luigi	Contributed to project design and manuscript revision through editorial feedback.	5%	5%

Acknowledgement

First and foremost, I would like to express my sincere gratitude to my supervisors Anthony Jakeman, Albert van Dijk, Siyuan Tian, and Luigi Renzullo for their dedication, guidance and support throughout my PhD program. Through their supervision, I developed solid scientific reasoning that will continue to shape my future research and professional career.

I am deeply grateful to Albert van Dijk for his intellectual generosity and guidance, which were instrumental in shaping my research direction and motivating my continued engagement with this topic beyond the thesis. His encouragement of curiosity, critical thinking and innovation, together with his emphasis on meaningful contributions to knowledge, has had a lasting influence on my understanding of the purpose and responsibility of scientific research. I would like to express my particular appreciation to Anthony Jakeman, whose continuous supervision, patience and encouragement sustained me through the later stages of my PhD. His confidence in my work, practical advice and steady support helped me navigate challenges and bring this thesis to completion.

I am sincerely thankful to Siyuan Tian and Luigi Renzullo for their guidance, constructive feedback and mentorship throughout my project. I am especially grateful for their support in organising internal reviews and facilitating access to key datasets and computational resources. I would like to thank Yi Liu and Peter Kanowski for invaluable supports during periods of uncertainty. Conversations with Yi consistently helped me regain confidence and clarity when I felt hesitant or discouraged. I sincerely appreciate Peter Kanowski for his kind support in navigating administrative matters at ANU, as well as for his ongoing care and encouragement around campus.

My sincere thanks also go to Tingbao Xu and Yun Chen for their valuable suggestions and encouragement on my research topic. I also thank Elisabeth Vogel for sharing historical simulations and ensemble forecast datasets that fundamentally supported this work. I am grateful to the ANU Institute for Water Futures (IWF), and in particular Barry Croke and Danlu Guo, for their insightful comments and support. Participation in Peter Hairsine's journal article writing workshop helped me understand the underlying logic of academic writing, which has been of lasting benefit.

I feel very fortunate to have shared my PhD journey with wonderful office mates, Takuya Iwanaga and James DelBene, whose companionship, encouragement and emotional support made daily research life far more manageable. I would also like to thank Takuya Iwanaga and Chad Burton for their warm assistance with coding and computing challenges, and for introducing me to the broader field of geospatial programming.

I am grateful to Ze Jiang and Shuang Liu from the UNSW Water Research Centre for sharing their work and for their encouragement and constructive suggestions. I also acknowledge the Australian Bureau of Meteorology (BoM) and its researchers for providing valuable advice and access to Australian Water Outlook project data essential to this research. I appreciate the technical support provided by the Australian National Computational Infrastructure (NCI).

My research was made possible through the support and resources of the Fenner School of Environment & Society at ANU. I also deeply appreciate the administrative and IT staff at Fenner, particularly Katie Liesinger, Harriet Torrens, Susan Ward and Winsome Law, for their warmth, patience and kindness. I am thankful to colleagues and fellow HDR students within Fenner for creating a stimulating and supportive research environment. I acknowledge Tony Boston, Wentao Ye, Moyang Liu, and Haiyan Liu for their friendship and support both within and beyond research.

I am extremely grateful to be part of the ANU Bushfire Research Centre of Excellence (BRCoE). Working with Marta Yebra, Nicholas Wilson and Li Zhao provided valuable opportunities to develop new skills and engage with external agencies. These experiences broadened my perspective on natural hazards research and have been both enjoyable and highly beneficial for my future career.

I am especially thankful to my close friends Xufei Luo, Haishan Li, Ting Wang, Yunong Zhou, and Tonghua Xie. Sharing the ups and downs of the PhD journey with laughter, conversations and mutual support made the process far less isolating.

Finally, I thank my family for their unconditional love and endless encouragement. Their unwavering support sustained me through the most challenging times. This thesis is dedicated to them, with deep gratitude and love.

Abstract

Drought ranks among the costliest natural hazards, exerting far-reaching consequences on natural ecosystems and human societies, particularly when severe soil moisture deficits induce vegetation deterioration (VD) stress. This stress impairs photosynthetic capacity and plant growth, triggering broader cascades—reduced crop yields, deteriorating pasture conditions, diminished terrestrial carbon stocks, and heightened wildfire. While hydro-meteorological variables such as soil moisture have long been recognised as key drivers of vegetation dynamics, operational drought early warning systems (DEWS) are often constrained by an emphasis on current conditions rather than predictive capabilities. Timely and skilful forecasts of vegetation drought impacts should enable more proactive and effective drought preparedness, management and mitigation. Predicting vegetation response to drought stress however remains challenging due to complex interactions between soil moisture deficits and plant health, compounded by varying response times and magnitudes across diverse ecosystems. To address this gap, this thesis advances the quantitative understanding of soil moisture–vegetation relationships across spatial and temporal scales, serving to develop a skilful early warning framework for multi-categorical vegetation deterioration risks.

The thesis encompasses three core research objectives: (i) examining the lower tail distribution of soil moisture and vegetation anomalies to identify critical thresholds for multi-level VD risk prediction; (ii) establishing and refining innovative prediction frameworks—specifically integrating a copula-based joint probability approach and a 'threshold event-matching' method—to capture both mild and severe VD events; and (iii) evaluating the extended skilful lead time and practical early warning horizon of the integrated framework in real-world drought risk management. By focusing on the growing season (May–October) and utilising soil moisture historical simulations and ensemble forecasts from the Australian Water Resources Assessment Landscape (AWRA-L) model, this thesis delivers a systematic approach for anticipating vegetation stress across large spatial domains.

Chapter 3 develops and tests a copula-based statistical framework, using simulated root-zone soil moisture anomalies to predict multiple severity levels of vegetation deterioration as observed in satellite-derived vegetation indices. Time-lagged correlation analyses indicate that mild vegetation deterioration events exhibit statistically meaningful predictive skill, whereas severe and infrequent events remain more difficult to anticipate, underscoring the need for enhanced modelling approaches. In Chapter 4, an innovative threshold event-matching approach is introduced, using in-situ soil moisture observations to capture rare but high-impact VD events. This method improves predictability, particularly in agricultural

regions, by pinpointing optimal soil moisture anomaly thresholds that correspond to substantial vegetation stress and providing lead times of approximately two to six weeks.

Building on these insights, Chapter 5 integrates the copula-based model and the threshold 'event-matching' framework. Using root-zone soil moisture simulations from the AWRA-L model over extensive spatial domains in southeastern Australia, this chapter demonstrates how each method's strengths can compensate for the other's limitations. Copula-based modelling tends to underestimate severe VD events, whereas the event-matching framework can overestimate mild–moderate categories; combined, however, they better capture the full spectrum of multi-categorical VD risks. The refined framework exhibits particular relevance for agricultural land use, offering more robust planning tools for drought-prone areas.

Subsequently, Chapter 6 applies the copula-based joint probability modelling and threshold event-matching approaches within an integrated framework to explore the use of ensemble soil moisture forecasts for VD risk prediction. The analysis examines probabilistic and deterministic forecasts at lead times ranging from one to six months and shows that forecast performance is spatially heterogeneous and generally declines with increasing lead time. Rather than demonstrating uniformly positive skill across the study region, the results illustrate how ensemble soil moisture forecasts can be structured to distinguish broader VD risk intervals, particularly at short lead times and during the late growing season. Case study analyses further show that forecasted probabilities of VD occurrence exceeding 0.5 tend to coincide with observed drought impacts in some regions, highlighting the potential of the framework to support impact-oriented drought risk awareness under favourable conditions. This extended early warning horizon is vital for decision-makers, stakeholders and farmers, enabling timely interventions such as crop selection, water allocation and landscape management to mitigate drought-related losses. Finally, Chapter 7 consolidates the findings and emphasises the framework's operational relevance for drought-prone agricultural systems. It outlines future research priorities, including enhancing spatial scalability and linking forecasts to economic cost–loss analyses.

Collectively, the thesis presents an operational predictive system that advances impact-based drought early warning. By integrating soil moisture anomalies with multi-categorical vegetation deterioration risks, it enables improved seasonal forecasting of drought impacts on vegetation. Through the application of ensemble soil moisture forecasts, the framework supports proactive decision-making and resource allocation for drought preparedness and mitigation. While its immediate application targets southeastern Australian agricultural systems, the methodologies, insights and modelling strategies developed here have broader relevance for other dryland regions worldwide.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Droughts generate complex cascading effects that originate from precipitation deficits, progress through soil moisture depletion and reduced runoff, and propagate through various stages of the hydrological cycle (Wilhite & Pulwarty, 2005; Mishra & Singh, 2010). Despite their slow-onset nature, droughts are costly and affect extensive geographical areas (Dai, 2011). Regarding drought effects on vegetated ecosystems specifically, severe soil moisture deficits directly stress plant biological functions, leading to agricultural disruptions including reduced crop yields, deteriorating pasture conditions for grazing, and volatile food market prices that may compromise food security (Farooq *et al.,* 2009; Rajsekhar *et al.,* 2015). These vegetation impacts extend to broader environmental consequences, such as decreased terrestrial carbon sequestration (Schwalm *et al.,* 2017), diminished ecosystem services (Van Oijen *et al.,* 2013), and increased wildfire risks (Jolly *et al.,* 2015). The World Meteorological Organization (WMO) and Global Water Partnership (GWP) have emphasised the advantages of impact-based Drought Early Warning Systems (DEWS), which transition drought management from reactive to proactive drought preparedness and mitigation planning by providing decision-makers with actionable intelligence at several months ahead (WMO & GWP, 2016).

Timely and skilful drought impact forecasting is crucial for delivering tangible information to stakeholders (Pozzi *et al.,* 2013; Pulwarty & Sivakumar, 2014), particularly for ecosystems located in water-limited areas. However, contemporary DEWS operating at regional, national and continental scales excel at monitoring current conditions, while their capacity to anticipate future vegetation situations with probabilistic confidence or scenario-specific performance remains limited. The satellite-derived Normalised Difference Vegetation Index (NDVI), computed from the differential reflection of near-infrared and visible radiation by photosynthetically active vegetation (Pettorelli *et al.,* 2016), serves as a widely adopted quantitative indicator of vegetation vigour and photosynthetic capability. This index enables systematic monitoring of vegetation health conditions across extensive spatial domains within DEWS frameworks (AghaKouchak *et al.,* 2015; Guo *et al.,* 2021). NDVI demonstrates strong correlations with multiple ecosystem values, including green biomass accumulation, agricultural productivity, and aspects of ecosystem functioning related to vegetation activity (Bai *et al.,* 2022; Huang *et al.,* 2019). Despite the availability of near-real-time NDVI observations through satellite platforms, the inherent retrospective nature of these measurements constrains their utility for prognostic applications in early warning systems. This limitation is particularly evident in current operational systems, which lack robust capabilities for forecasting drought-induced deterioration of vegetation greenness conditions.

While the slow-onset characteristics of drought theoretically provide adequate lead time for intervention, developing reliable prediction frameworks for relationships between hydro-meteorological variables and vegetation greenness remains constrained by significant biophysical and statistical challenges. These challenges encompass complex vegetation physiological responses across diverse ecosystems, including intricate response chains between drought conditions and vegetation behaviour, and variable time lags between drought event onset and observable impacts (Anderegg *et al.,* 2020; Yuan *et al.,* 2019). The temporal dynamics of these soil moisture-vegetation interactions exhibit notable complexity, characterised by non-linear responses and ecosystem-specific lag effects that can range from weeks to months (Konings *et al.,* 2021; Ukkola *et al.,* 2016). Non-linear response mechanisms within ecohydrological systems can amplify input biases from climate forcings, which traditional bias correction methods may not fully address (Johnson et al., 2024). The statistical limitations in modelling these relationships primarily stem from the relatively small sample size of extreme drought events and critical gaps in historical drought impact records, with reliable satellite-based assessment only available since the early 2000s (Dorigo *et al.,* 2017; Vogel *et al.,* 2021).

Previous research has demonstrated that root-zone soil moisture availability is the dominant driver of vegetation dynamics in dryland ecosystems (Zhang *et al.,* 2023; Feng & Fu, 2016). In the Australian context, soil moisture-vegetation coupling is particularly sensitive to soil moisture deficits and their timing across diverse ecological and agricultural systems (Ukkola *et al.,* 2016; De Kauwe *et al.,* 2020), especially in the sheep-wheat belt of southeastern Australia (Chenu *et al.,* 2018). The region encompasses a variety of climate regimes, including the temperate southeast, cold semi-arid inland, cool-humid eastern uplands, and Mediterranean west coastal areas (Gallant *et al.,* 2012). These diverse climates, characterised by pronounced rainfall seasonality, high interannual variability, and recurrent drought episodes, create a complex hydro-ecological environment where soil moisture dynamics play a deterministic role in shaping vegetation productivity and agricultural outcomes (Akuraju *et al.,* 2017; van Dijk *et al.,* 2013).

The Australian Water Resources Assessment Landscape (AWRA-L) model is widely recognised for providing reliable continental simulations of root-zone soil moisture (Frost *et al.,* 2018). It underpins the Australian Water Outlook operated by the Bureau of Meteorology, providing near-real-time volumetric soil water content (m^3 m^{-3}) at a 5 km spatial resolution from 1911 to the present (Jones *et al.,* 2020), together with operational soil moisture outlooks available from 2019 with skilful lead times of up to three months (Vogel *et al.,* 2021). Although AWRA-L outputs effectively capture soil moisture conditions and anomalies, they do not directly predict future vegetation growth. Therefore, by quantifying the relationship between soil moisture deficits and vegetation deterioration, it is possible to develop a predictive framework to forecast vegetation status. Such predictions would equip stakeholders and decision-makers with clearer insights into how forecast root-zone soil moisture conditions may affect the growth of natural vegetation, pastures and crops in upcoming seasons. Given these knowledge gaps, there exists a critical need to advance vegetation impact-based early warning capabilities in dryland ecosystems through enhanced forecasting accuracy at operationally relevant lead times. This research specifically emphasises southeastern Australia's vulnerable vegetation systems to strengthen early warning systems for proactive drought risk management.

1.2. Research aims and objectives

The overall aim of this thesis is to advance the quantitative understanding of soil moisture–vegetation relationships across spatial and temporal scales, with the dual goals of improving theoretical insight into drought–vegetation dynamics and informing the development of early warning systems for vegetation deterioration risks. More specifically, it investigates how model-generated soil moisture simulations and ensemble forecasts can support both scientific understanding and predictive applications at scale. Focusing on the southeastern Australian wheat-sheep belt during the growing season (May-October), the specific objectives of this research were:

1. Characterising soil moisture–vegetation interactions: Examine the temporal and statistical relationships between extreme soil moisture deficits and vegetation stress, with a focus on lower-tail distributions and spatial patterns of vegetation anomalies. This involves evaluating copula-based joint probability models and “threshold event-matching framework” that leverage state-of-the-art model-simulated soil moisture to represent multi-categorical vegetation deterioration risks, ranging from mild to exceptional severity levels across both native and agricultural landscapes.

2. Developing a process-informed predictive framework: design and validate a modelling approach that optimises soil moisture thresholds to predict varying vegetation deterioration severity levels, integrating copula-based modelling and threshold event-matching framework to capture both frequent and uncommon

vegetation stress events across diverse land use types. The objective focuses on improving predictive accuracy and interpretability by embedding event-based process understanding, thereby enhancing the model's applicability across large spatial scales.

3. Assessing predictive performance and practical relevance: evaluate the operational utility of this integrated framework by assessing its skill and lead time using ensemble soil moisture forecasts. This aims to translate model-based insights into actionable intelligence for early warning practices in Australia's dryland ecosystems.

The thesis is structured around four key research questions (RQs) below, each aligned with a core analytical chapter (Chapters 3 to 6). These RQs collectively support the overarching research objectives (Objectives 1-3) outlined in Subsection 1.2:

- RQ1 (Chapter 3): *Can the joint probability distribution approach (e.g., copula-based models) be used to develop a statistical framework for predicting multi-categorical vegetation deterioration events based on historical root-zone soil moisture simulations?*

- RQ2 (Chapter 4): *Do in-site soil moisture observations, through the 'threshold event-matching' approach, offer potential capability in predicting infrequent-severe vegetation deterioration events, and if so, what is the lead time?*

- RQ3 (Chapter 5): *Can the 'threshold event-matching' approach improve predictions of vegetation deterioration severity at large scales, and how does integrating it with copula-based models enhance prediction accuracy for different land use types in southeastern Australia?*

- RQ4 (Chapter 6): *Up to what lead time can ensemble soil moisture forecasts skilfully predict vegetation deterioration risks, and how can deterministic and probabilistic forecasts be applied into drought early warning systems?*

1.3. Thesis structure

This thesis comprises four key research chapters (Chapters 3–6) and one literature review chapter (Chapter 2). The chapters are interconnected, each building on the previous to present a cohesive framework for integrating model-generated soil moisture simulations and ensemble forecasts into a skilful prediction system for multi-categorical vegetation deterioration risks across large spatial scales (Fig. 1.1). The thesis follows a logical progression, moving from the establishment of statistical relationships (Chapters 3 and 4), to the integration and refinement of prediction frameworks (Chapter 5), and ultimately applying forecast products in practical early warning contexts (Chapter 6).

In addition to the thematic progression, the chapters also differ in their temporal prediction scope. Chapters 3 to 5 focus on the inherent lagged response between soil moisture conditions and vegetation dynamics under near-real-time conditions, whether derived from model-simulated or in-situ observed soil moisture. These chapters aim to characterise and utilise the short-term water stress signals driven by vegetation–soil moisture coupling. In contrast, Chapter 6 investigates the extent to which seasonal ensemble soil moisture forecasts can extend this temporal prediction window. The total forecast horizon in the integrated framework thus comprises two components: (i) the vegetation response time (d) and (ii) the forecast lead time (L), as conceptually illustrated in Fig. 1.2. Each chapter is outlined as follows:

Chapter 1 introduces the research background, identifies key knowledge gaps, and outlines the overall research aims, key research questions, and specific objectives. Chapter 2 provides an in-depth review of existing approaches to forecasting drought impacts, including those that incorporate atmospheric, oceanic and hydro-meteorological variables. Particular attention is given to satellite-based remote sensing methods for large-scale vegetation condition observation, which form the basis for assessing vegetation drought impacts in this thesis. It also evaluates current operational drought early warning systems (DEWS) and land surface hydrological forecasting products globally, and synthesises recent modelling studies on soil moisture forecasting. This literature review identifies the knowledge gaps and research scope of the thesis, and suggests future research directions for improving drought impact prediction on vegetation.

Chapter 3 investigates the relationship between vegetation dynamics and soil moisture variability in southeastern Australia, developing a prediction framework using copula-based joint probability modelling. This modelling approach has been widely applied in fields such climate risk analysis, hydrology and ecohydrology, but is here adapted to predict multi-categorical vegetation deterioration severity. The chapter quantifies the relationship between root-zone soil moisture and vegetation anomalies through time-lagged Spearman correlation and assesses the predictive performance of the framework for multi-categorical vegetation deterioration (VD) risks. It highlights the copula-based model's ability to predict mild vegetation drought impacts, while identifying its limitations for infrequent severe events.

Chapter 4 introduces an innovative *threshold event-matching* framework to enhance the prediction of infrequent-severe VD events, using in-situ soil moisture observations from 93 sites across the central and western United States. The approach significantly improves prediction accuracy for rare, high-impact VD events. The chapter evaluates the effectiveness of soil moisture observations in predicting satellite-derived vegetation drought impacts, demonstrating enhanced predictability, especially for agricultural regions.

Chapter 5 extends the *threshold event-matching* framework to large-scale soil moisture simulations over southeastern Australia. The chapter refines the prediction framework by integrating copula-based modelling with the event-matching approach, enabling the capture of both frequent mild and infrequent severe VD events across southeastern Australia. It evaluates the performance of the optimal integrated framework in predicting VD risks, highlighting the complementary strengths of copula-based and event-matching models in various vegetation systems.

Chapter 6 demonstrates the operational applicability of the integrated framework in southeastern Australia by applying it to ensemble soil moisture hindcasts (i.e., retrospective forecasts) datasets at lead times ranging from 1 to 6 months. The chapter assesses the predictive skill and practical early warning horizons of probabilistic and deterministic VD forecasts at varying lead times. It identifies skilful lead times for VD risk prediction, particularly for mild to severe VD categories, and shows that the framework provides reliable early warning signals for drought-induced vegetation stress, especially during the late growing season in agricultural zones.

Chapter 7 concludes the thesis, summarising the research findings, addressing the knowledge gaps, and proposing avenues for future research. It also discusses the limitations of the methodology and suggests how the early warning framework can be further improved for better decision-making in agricultural drought risk management.

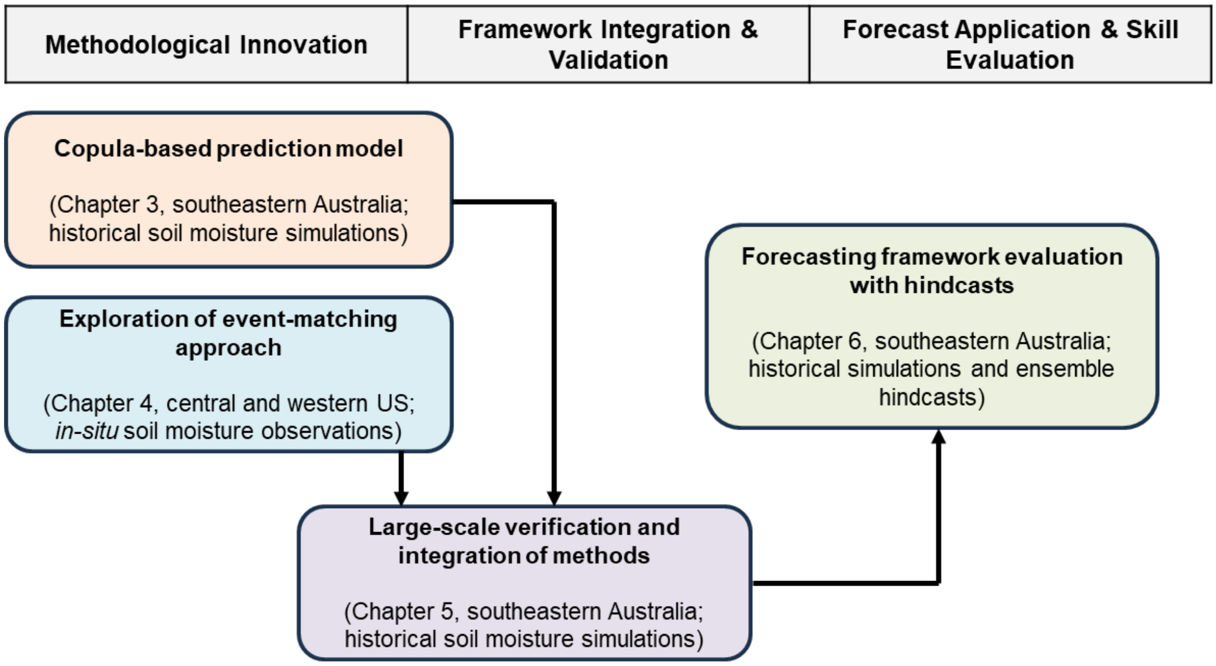


Fig. 1.1. Schematic overview of the thesis structure and methodological progression across Chapters 3–6.

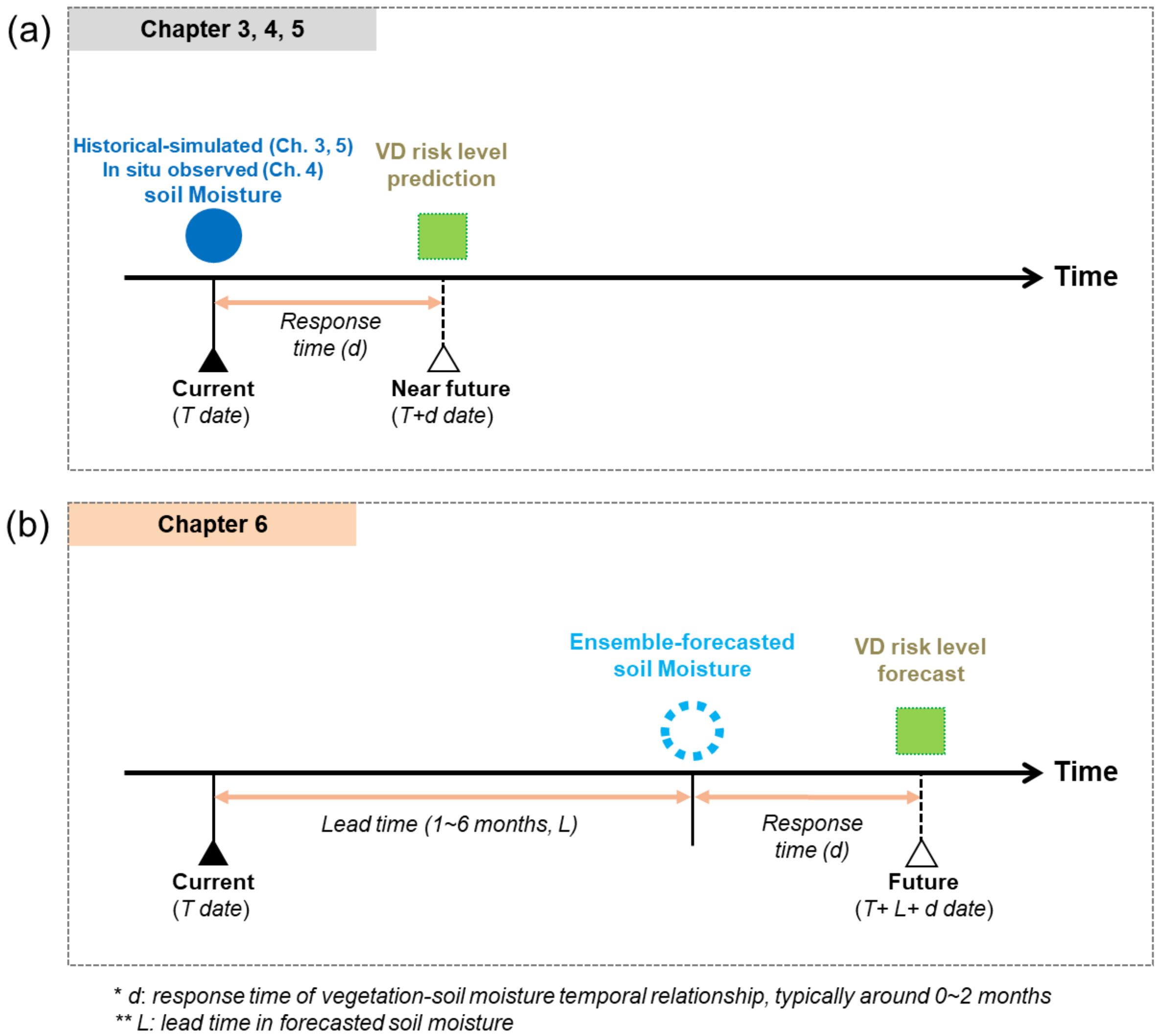


Fig. 1.2. Conceptual illustration of focused temporal horizons across Chapters 3–6. (a) Near-future prediction using model-simulated or in-situ observed soil moisture (Chapters 3–5); (b) Extended lead-time prediction using ensemble-forecasted soil moisture (Chapter 6). The total forecast horizon combines vegetation response time (*d*) and soil moisture forecast lead time (*L*).

Chapter 2

Integrating vegetation impact forecasting into drought early warning systems: a comprehensive review

2.1. Drought as a climatic hazard: concepts and definitions

2.1.1. Drought types and classification

Drought is a widespread climatic extreme that affects nearly half of the Earth's terrestrial surfaces, exerting significant impacts on agriculture, food security, hydropower, industry, health, and social stability (WMO, 2006). These impacts are anticipated to intensify in the twenty-first century, driven by climate change (Dai, 2013; Schwalm *et al.,* 2017). Unlike other climatic hazards, such as floods, drought lacks a precise and universally accepted definition due to its complex causal mechanisms, varying duration, intensity, spatial extent, and associated impacts (Mishra and Singh, 2010). Drought definitions are commonly classified into four types (Fig. 2.1), based on underlying causal mechanisms and resultant impacts (Wilhite and Glantz, 1985; Mishra and Singh, 2010; Wilhite, 2005, 2016):

- Meteorological drought: A prolonged period (months or years) characterised by precipitation levels significantly below the regional long-term average.
- Agricultural drought: A condition where insufficient soil moisture impairs plant growth, leading to reductions in crop yield and pasture productivity.
- Hydrological drought: A deficiency in precipitation that negatively affects surface or groundwater resources (e.g., streamflow, groundwater levels, reservoirs, and lakes) relative to long-term averages.
- Socio-economic drought: Occurs when the supply of water resources and water-dependent economic activities fails to meet societal demand.

Although these definitions differ, they are inherently interconnected and cannot be viewed in isolation (Wilhite and Glantz, 1985). In practice, classifying a drought event into a singular category proves challenging and may vary depending on the specific objectives of the study (Mishra and Singh, 2010). For example, a period of decreased precipitation may initiate meteorological drought, which through continued evaporation, can evolve into agricultural drought. If diminished plant function leads to a significant decline in crop yields, agricultural drought may subsequently escalate into socio-economic drought.

A common source of confusion arises between the terms “drought,” “aridity,” and “dry conditions.” Unlike aridity, which is a permanent climatic feature of low rainfall regions, drought is a transient phenomenon that can manifest in both arid and humid regions (Wilhite, 2005). Additionally, droughts are statistically infrequent events, typically occurring once every twenty-five years, whereas dry conditions and water-related stress can arise multiple times per year (Sun *et al.,* 2023).

Several studies have misinterpreted the term "agricultural drought" by equating it solely with soil moisture anomalies without considering vegetation conditions, particularly in energy-limited regions such as higher latitudes (van Hateren *et al.,* 2021). Agricultural drought is generally considered to occur when soil moisture deficits become anomalously severe relative to typical seasonal conditions and heavily impair plant growth and crop yields (Sheffield and Wood, 2011; van Dijk *et al.,* 2013). This creates a practical challenge in defining the onset of agricultural drought events. Soil moisture within the plant root zone is essential for vegetation health and agriculture. While the actual rooting depths vary substantially across plant species and ecosystems, the large-scale land surface models typically represent this as the upper 0-1m of the soil layer (Yang *et al.,* 2016).

To address this issue clearly, some researchers advocate distinguishing explicitly between "soil moisture drought," reflecting soil water deficits, and "vegetation drought," focusing on impaired plant growth (Cunha *et al.,* 2015; Weng *et al.,* 2023; Zeng *et al.,* 2023; W. Li et a., 2023). Furthermore, the term "ecological drought" has been proposed to broadly encompass drought impacts on both agricultural and natural ecosystems, including vegetation responses and effects on the carbon cycle (Crausbay *et al.,* 2017, 2020; Bradford *et al.,* 2020).

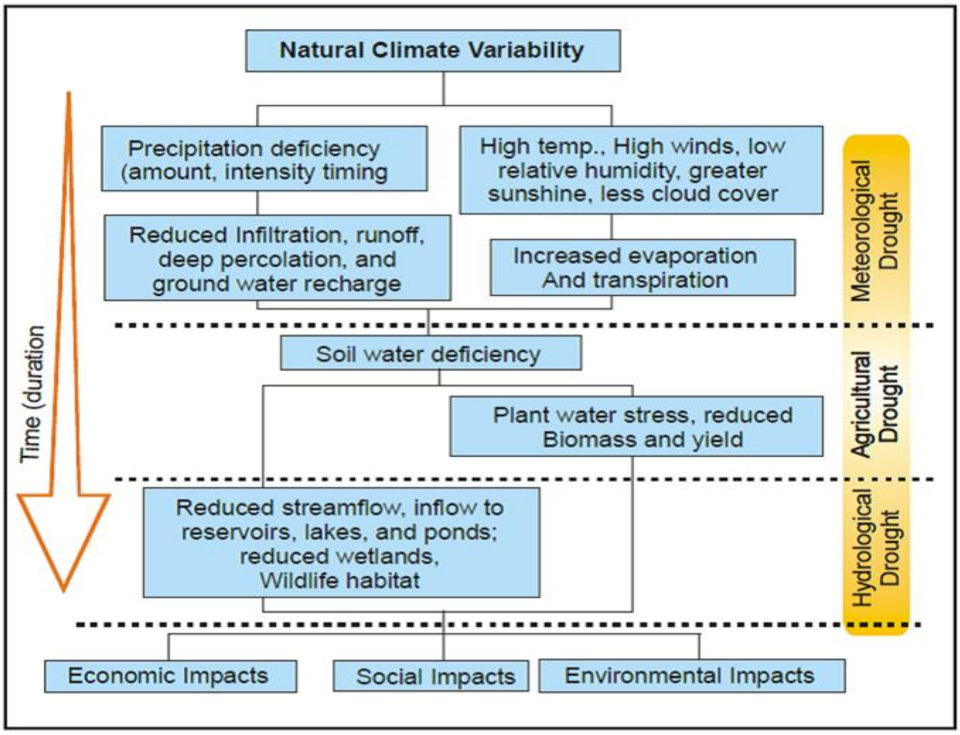


Fig. 2.1. Different types of droughts and their interactions (Adapted after Van Loon, 2015)

2.1.2. Hydro-meteorological indicators for drought analysis

Hydro-meteorological indices and indicators provide quantitative or categorical assessments of drought characteristics such as onset, severity, frequency and duration (Mishra and Singh, 2010). Indicators describe drought conditions through variables like precipitation (mm), air temperature (°C), streamflow (m^3/s), reservoir storage (m^3) and soil moisture (m^3/m^3), whereas indices are numerical representations of drought intensity derived from these variables (WMO and GWP, 2016). These indices are commonly mapped across regions to characterise the spatial and temporal evolution of drought (Farahmand and AghaKouchak, 2015). Existing approaches can be broadly grouped into simple, multiple-indices composited, and categorical hybrid approaches.

Simple approach derives meteorological or hydrological indices directly from a single variable, typically using long-term historical observation records or model simulations (Hao *et al.,* 2017). Widely used indices include the Standardized Precipitation Index (SPI; McKee *et al.,* 1993) and the Standardized Precipitation Evapotranspiration Index (SPEI; Vicente-Serrano *et al.,* 2010), both computed over multiple accumulation periods (e.g., 1, 3, 6 months). The SPI, recommended by the World Meteorological Organization (WMO) for global drought monitoring (WMO, 2009), is valued for its simplicity and comparability across climates. However, its reliance on precipitation alone limits its ability to represent drought processes in regions where evapotranspiration and soil moisture dynamics play a dominant role (Vicente-Serrano *et al.,* 2010). SPEI partially addresses this limitation by incorporating temperature-driven evaporative demand, but its performance remains sensitive to the choice of PET formulation and does not explicitly represent root-zone soil moisture constraints on vegetation (Hao & AghaKouchak, 2013).

Multiple-indices composited approach aim to capture the multi-dimensional nature of drought by integrating meteorological, hydrological and vegetation-related variables (Mishra & Singh, 2010). Indices such as the Vegetation Health Index (VHI), defined as VHI = αVCI + (1-α)TCI, combines information on Vegetation Condition Index (VCI) and thermal stress represented by Temperature Condition Index (TCI) (Brown *et al.,* 2013; Kogan, 1995). Nevertheless, their interpretation can be complicated by subjective weighting schemes, differing sensitivities of component variables, and limited transferability across regions with contrasting climate–vegetation regimes.

Categorical hybrid approach integrate multiple physical indicators with expert assessment to characterise drought severity and impacts (WMO, 2006). For example, the United States Drought Monitor (USDM) classifies drought into multiple severity categories (D0–D4) using diverse sources of physical indicators and on-the-ground reports of drought impacts evaluated by experts' interpretation (Svoboda *et al.,* 2002; Z. Li *et al.,* 2023). Similarly, the Combined Drought Indicator (CDI) of the European Drought Observatory (EDO) integrates the 3-month SPI, model-simulated soil moisture, and satellite-based fraction of Absorbed Photosynthetically Active Radiation (fAPAR) into three drought levels: watch, warning, and alert (Sepulcre-Canto et al. 2012). While such systems are highly effective for operational monitoring and communication, their reliance on expert judgement and region-specific calibration can limit reproducibility and applicability outside their original domains.

Although a wide range of drought indices and indicators has been developed, a universally accepted drought metric remains elusive due to the context-specific nature of drought processes and impacts (Hao & Singh, 2015; Hayes *et al.,* 2011; Mishra & Singh, 2010; Zargar *et al.,* 2011). In this thesis, I prioritise soil moisture-based indicators derived from land surface models and satellite observations because they provide a physically interpretable link between hydro-meteorological forcing and vegetation stress, are directly relevant to agricultural and ecological drought, and offer spatially continuous, repeatable coverage suitable for early warning applications (Sheffield *et al.,* 2012). This focus aligns with the objective of characterising drought-induced vegetation deterioration rather than drought conditions defined solely by atmospheric anomalies.

2.1.3. Thresholds for determining drought event onset and severity

Drought severity is commonly assessed by the degree of departure of drought indices from their climatological normal values (Svoboda *et al.,* 2002; Mishra and Singh, 2010). In practice, drought onset and severity are identified using predefined thresholds; however, considerable inconsistency remains in how these thresholds are selected and applied across studies (Singh *et al.,* 2016; Bachmair *et al.,* 2016). Standardised indices such as the SPI and the SPEI are typically interpreted using fixed thresholds of standard deviations from the mean. For example, drought conditions are often defined when SPI or SPEI falls below −1, with severity further classified as moderate ($-1.5 < \text{SPI} < -1$), severe ($-2 < \text{SPI} \leq -1.5$), and extreme ($\text{SPI} \leq -2$) (McKee *et al.,* 1993; WMO, 2009; Hao and Singh, 2015). Although expressed in standardised units, these thresholds correspond to specific percentiles of the historical record, as SPI and SPEI are derived by mapping climate variables onto a standard normal distribution.

Similarly, percentile-based thresholds (e.g., the 5th or 10th percentile) are widely used for hydrological indicators such as streamflow, reservoir storage and groundwater levels (Van Loon & Van Lanen, 2012; Wanders *et al.,* 2017). Information derived from these drought indices and thresholds underpins operational drought early warning systems (DEWS), such as the USDM–based drought outlooks guiding drought risk assessment, official drought declarations, and management interventions (Wilhite *et al.,* 2014; Hao *et al.,* 2018).

2.2. Remote sensing-based assessment of drought impacts on vegetation

2.2.1. Physiological responses of vegetation to water stress

Agricultural drought has immediate and substantial impacts on vegetation, significantly affecting ecosystems, agricultural production, and human welfare (Wilhite, 2000; Mishra and Singh, 2010; van Dijk *et al.,* 2013). These impacts include reduced vegetation greenness and productivity, increased tree mortality, altered ecosystem carbon cycling, and elevated bushfire risks (Zhao and Running, 2010; van der Molen *et al.,* 2011; Schwalm *et al.,* 2017). However, unlike hydrological drought, which can be directly quantified using water levels or flow metrics, identifying agricultural drought through vegetation responses poses practical challenges due to varying vegetation physiological responses to similar drought intensities across ecosystems and climate conditions (Zhao and Running, 2010; van der Molen *et al.,* 2011).

The physiological mechanisms underlying vegetation responses to soil water stress are illustrated in Fig 2.2. Persistent reductions in soil water availability negatively affect critical physiological processes, including photosynthesis, transpiration and nutrient uptake, consequently impairing vegetation growth and productivity (Tyree, 1999; Nobel, 2009; Damm *et al.,* 2018). Under short-term soil moisture deficits, vegetation typically responds by closing stomata to reduce transpiration, which leads to elevated canopy temperature (Maselli *et al.,* 2009). When soil moisture deficits become anomalously severe and prolonged relative to typical seasonal conditions, photosynthetic capacity declines further, chlorophyll content decreases, morphological changes are included, and plant senescence may occur (Inoue *et al.,* 1990; Nobel, 2009). These progressive responses describe vegetation behaviour across increasing levels of water stress, providing a physiological basis for understanding vegetation impacts under drought conditions.

Vegetation physiological response to drought is nonlinear and species-dependent, influenced by structural traits, phenological stages, and drought characteristics (Farooq *et al.,* 2009; Wu *et al.,* 2021). Plants with adaptations such as deeper roots can access subsurface water reserves (Farooq *et al.,* 2009). Research confirms nonlinear relationships between vegetation productivity and soil moisture deficits globally (Li *et al.,* 2023), with impacts varying according to ecosystem-specific thresholds (Guha *et al.,* 2021). The temporal complexity of vegetation response, including lagged effects dependent on drought progression and plant phenology (Vicente-Serrano *et al.,* 2013), complicates reliable identification of agricultural drought onset and severity from vegetation conditions.

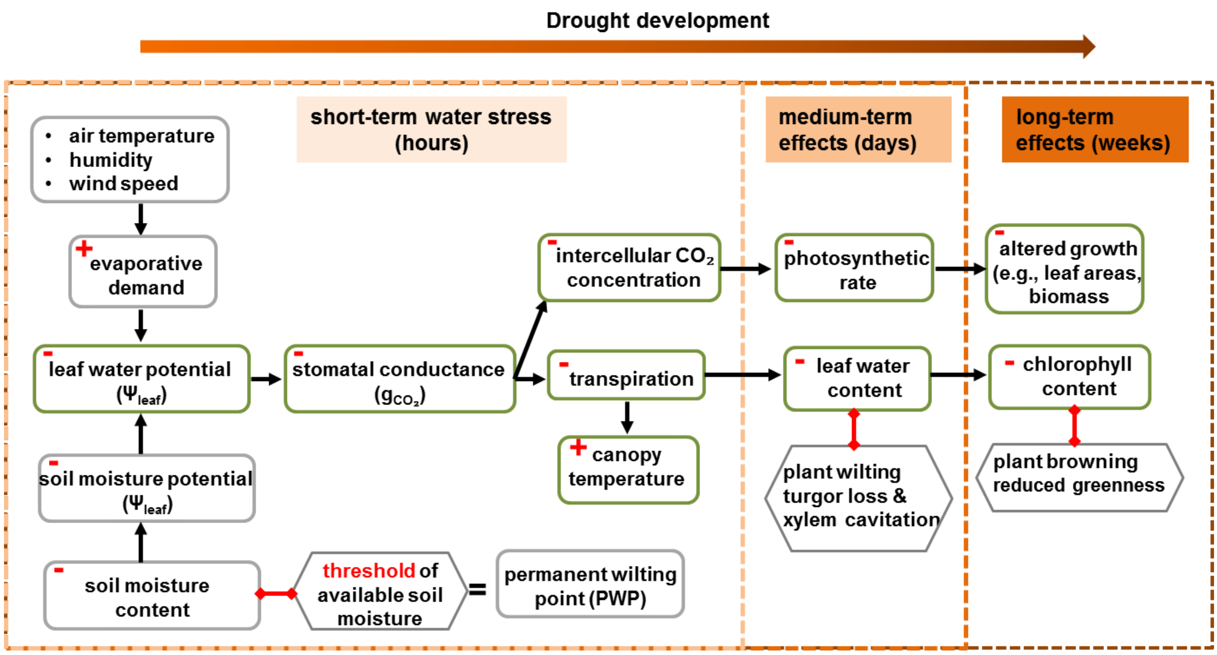


Fig. 2.2. Schematic representation of the sequential physiological processes involved in vegetation response to water stress, spanning short-term (hours), medium-term (days), and long-term (weeks) timescales. Red symbols "+" and "–" denote positive (stimulatory) and negative (inhibitory) effects, respectively, indicating whether a process enhances or suppresses another.

2.2.2. Satellite-derived indicators of vegetation water stress

Satellite-based remote sensing has become a key tool for monitoring vegetation drought impacts at large scales, overcoming the challenges of ground-based station measurements, such as cost and timeliness (Konga, 1997). Over the past three decades, remote sensing has enabled spatially continuous vegetation assessments conducted at regular intervals, significantly improving the monitoring of drought-related vegetation stress at both regional and global scales (AghaKouchak *et al.,* 2015; Krishnamurthy *et al.,* 2018).

While short-term drought responses like stomatal closure and changes in transpiration are difficult to detect remotely (Sheffield and Wood, 2011; Damm *et al.,* 2018), long-term soil moisture deficits can be monitored through satellite observations. These deficits lead to changes in plant pigment composition and canopy structure, affecting radiation absorption and scattering, which can be captured through remote sensing (Konga, 1997; Maselli *et al.,* 2009; Anyamba and Tucker, 2012). Key vegetation parameters derived from satellite observations include photosynthetic activity, canopy greenness, leaf area, canopy temperature, canopy water content, canopy structure, biomass, and transpiration rates (AghaKouchak *et al.,* 2015; Damm *et al.,* 2018; W. Li *et al.,* 2023). Variables related to plant water status, such as leaf water potential, are not directly observed by remote sensing but may be indirectly inferred through proxies such as microwave-based vegetation optical depth (VOD) that are sensitive to canopy water content and biomass (Zhang *et al.,* 2019).

The following sections outline remote sensing approaches that span visible to microwave wavelengths for detecting and assessing vegetation drought impacts (Fig. 2.3).

- Canopy Greenness: Canopy greenness serves as a key indicator (proxy) of vegetation health and photosynthetic capacity (Tucker & Choudhury, 1987; Sheffield *et al.,* 2012). Vegetation indices (VIs), such as the Normalized Difference Vegetation Index (NDVI), which quantifies the difference between near-infrared (NIR) and red reflectance, are widely employed to assess greenness (Rouse *et al.,* 1974). NDVI is calculated using the formula: NDVI = (ρNIR – ρRed) / (ρNIR + ρRed) (Tucker, 1979), where ρ denotes surface reflectance in the NIR and Red spectral bands, respectively. The Enhanced Vegetation Index (EVI) enhances sensitivity, particularly in dense canopies, by mitigating spectral noise from soil and atmospheric interference (Huete *et al.,* 2002).
- Canopy temperature and evapotranspiration: Land Surface Temperature (LST), derived from thermal infrared data, offers insights into canopy temperature and drought-induced water stress (Anderson *et al.,* 2013). The Temperature Condition Index (TCI) helps assess temperature-related vegetation stress (Kogan, 1995a). The ECOSTRESS (ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station) mission provides valuable thermal infrared data that can be used to directly measure plant canopy temperature, which is an indicator of vegetation water stress and evaporative cooling (Fishier *et al.,* 202; González-Dugo *et al.,* 2021). One of the main ECOSTRESS products is the Evaporative Stress Index (ESI), which quantifies evaporative stress in plants by comparing their actual temperatures to (potential) temperatures expected if they were not stressed through the water and energy budget model (i.e., Atmosphere-Land Exchange Inverse (ALEXI) model (Anderson *et al.,* 2021)).
- Vegetation water content: Indices like the Normalized Difference Infrared Index (NDII), Normalized Difference Water Index (NDWI) and live fuel moisture content (LFMC) use spectral bands (e.g., shortwave infrared) sensitive to water content to monitor vegetation water status (Caccamo *et al.,* 2011; Gao, 1996; Yebra *et al.,* 2013).

- Vegetation photosynthetic activity: Solar-induced chlorophyll fluorescence (SIF) has gained popularity for estimating gross primary productivity (GPP) dynamics (Frankenberg *et al.,* 2011; Wood *et al.,* 2017; Smith *et al.,* 2018). SIF arises from chlorophyll fluorescence emitted by plant leaves in the red and near-infrared spectral regions following the absorption of solar radiation, providing a direct physiological link to photosynthetic activity and showing strong correlations with ground-measured GPP across diverse vegetation types (Lee *et al.,* 2013; Yoshida *et al.,* 2015; Song *et al.,* 2018). SIF responds rapidly to water and heat stress, reflecting actual photosynthetic activity (Wang *et al.,* 2019). The NIRv index, which combines NDVI and NIR reflectance, is closely linked to SIF and GPP, providing a reliable measure of vegetation photosynthetic potential (Badgley *et al.,* 2019; Zhang *et al.,* 2022).

- Vegetation aboveground biomass: Microwave sensors, such as resulting brightness temperature (Tb) from long-wavelength bands retrievals, can be used to estimate vegetation optical depth (VOD) (Liu *et al.,* 2007, 2009, 2011). VOD provides a frequency-dependent metric of vegetation aboveground biomass and total water content changes and has also been used to infer plant water potential (Konings and Gentine, 2016).
- Plant structure: Satellite-based remote sensing techniques, such as Light Detection and Ranging (LiDAR), Synthetic Aperture Radar (SAR), and Interferometric Synthetic Aperture Radar (InSAR), offer unique capabilities to assess vegetation structure and changes induced by drought stress, including vegetation height and leaf angle (Kukenbrink *et al.,* 2017).

In summary, strategies for quantitatively assessing drought-related vegetation stress at large scale can be broadly categorised into two approaches:

(i) Satellite observations alone, which rely on measurements of plant properties or processes influenced by variations in water availability. These include indicators of canopy greenness, photosynthetic activity, land surface temperature, plant water status, and canopy structure; and

(ii) Simulated bio-physical variables, such as crop yields and pasture growth, derived from models that use plant-related parameters (e.g., greenness) together with environmental drivers, including rainfall, air temperature, vapour pressure deficit (VPD), and irradiance.

These parameters are used to drive process-based models, which simulate vegetation productivity metrics (e.g., fAPAR, Leaf Area Index (LAI), GPP), aboveground biomass, and transpiration rates (e.g., the ratio of ET to PET). These model outputs can then be applied to evaluate the impacts of drought on vegetation health (Reichstein *et al.,* 2007; Vicca *et al.,* 2016; Yu *et al.,* 2017; Damm *et al.,* 2018; Ahmadi *et al.,* 2019).

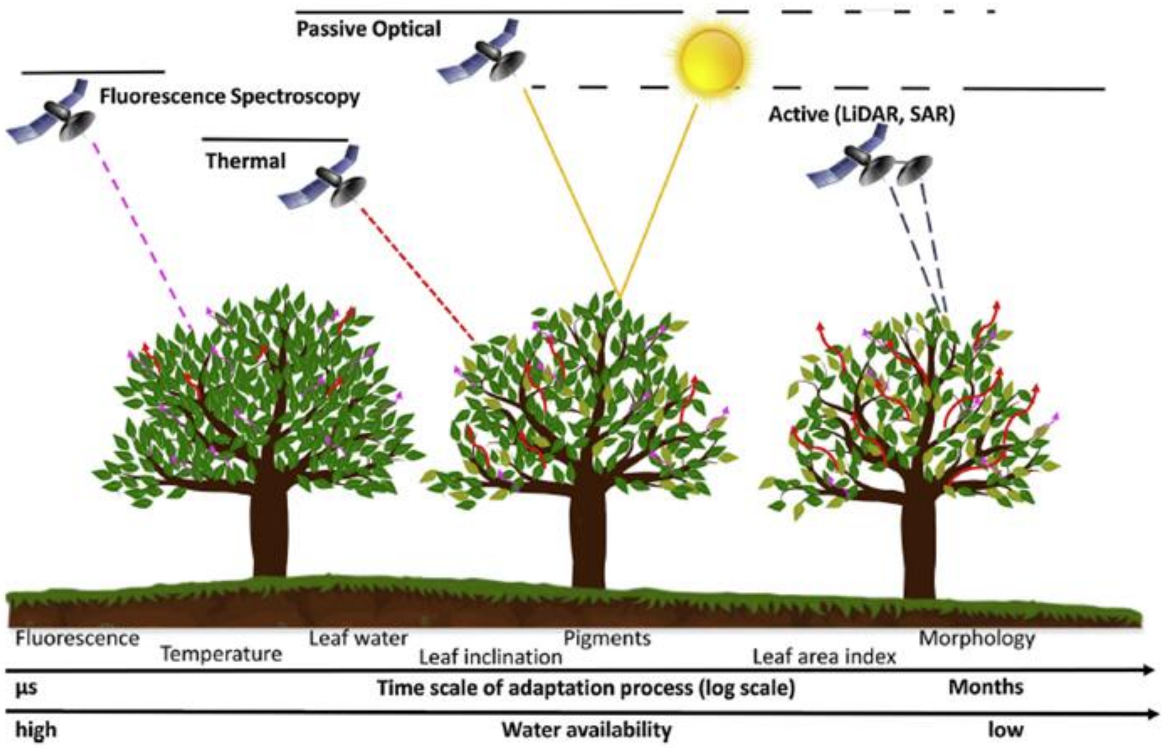


Fig. 2.3. Recent advances in remote sensing of plant-water relations (Damm *et al.,* 2018)

2.3. Drought early warning systems (DEWS): principles and components

2.3.1. Definition and importance of DEWS

Early Warning Systems (EWS) have been implemented globally to monitor, forecast, and alert communities about various natural hazards, including hydro-meteorological events (e.g., floods, droughts), geological hazards (e.g., earthquakes, tsunamis), and biological threats (e.g., epidemics, infestations) (UN-ISDR, 2011; UNEP, 2012). The core principle of EWS is to provide accurate, timely predictions of both short- and long-term risks, thereby enhancing the ability to manage and mitigate their societal, economic and environmental impacts (UNEP, 2012).

Recognising the need for a shift from in-drought reactive to proactive disaster management, many regions have prioritised drought preparedness and climate resilience (UNEP, 2012; FAO, 2019). Reliable drought early warning systems (DEWS) are essential for this transformation, relying on integrated risk assessments grounded in scientific data (Dai, 2011; Hayes *et al.,* 2012). A comprehensive DEWS involves three key elements (Fig. 2.4): (i) risk knowledge via technical monitoring and forecasting; (ii) effective communication of warnings;

and (iii) the capacity for mitigation and preparedness within communities and governments (WMO & GWP, 2016). Failure in any of these components undermines the effectiveness of the entire system (UNEP, 2012). With timely information from DEWS, water and land managers, policymakers and the public can take proactive measures to reduce vulnerability and reap significant benefits (UNEP, 2012; FAO, 2019). However, the definition of "early" in warning delivery is context-dependent, shaped by political, socio-economic and cultural factors (UNEP, 2012; Hughes *et al.,* 2022).

Fig. 2.4. Three pillars of integrated drought management (WMO & GWP, 2016)

2.3.2. Monitoring and forecasting components in DEWS

Drought Early Warning Systems (DEWS) track and assess climatic, hydrological, and water supply conditions to deliver timely information before, during, or at the onset of drought, prompting action in drought risk management and preparedness (WMO, 2006; WMO and GWP, 2016). These systems ideally combine both monitoring and forecasting components, providing near-real-time tracking and early notifications with lead times from a few months to several seasons (WMO, 2006; UNEP, 2012). The operational details of DEWS are further discussed in Subsection 2.3.3.

2.3.2.1 Monitoring component

The monitoring component tracks drought propagation, severity and spatial extent, helping governments decide when to officially declare a drought (FAO, 2019). It employs a comprehensive approach to monitor drought development, addressing all types of drought, not just specific ones like hydrological drought. Key parameters such as monthly precipitation volume are monitored to determine drought onset, end and spatial characteristics (WMO, 2006). Advances in technology enable real-time monitoring of drought phenomena (e.g., the status of satellite-derived vegetation health index) across both space and time. Drought severity is assessed using indices on weekly, monthly, seasonal and annual scales, with results issued as drought maps. These products support situational awareness, inform drought planning processes, and contribute to impact assessment and response coordination, including applications in agricultural management and risk communication (Pulwarty and Sivakumar, 2014; Senay *et al.,* 2015).

2.3.2.2 Forecasting component

The forecasting component uses climate and atmospheric information, including historical climate records, real-time meteorological observations, and numerical weather forecasts, to estimate likely future climate and water supply conditions (WMO & GWP, 2016). Forecasts are visualised in drought outlook maps showing predicted precipitation and temperature varying trends up to six-months (often three months) in advance (WMO, 2021; Charles *et al.,* 2020). Several operational hydrological prediction systems have been developed to provide meteorological and hydrological drought forecasts, which rely on seasonal predictions from major numerical weather prediction centres such as the National Oceanic and Atmospheric Administration's (NOAA's) National Centers for Environmental Prediction (NCEP) and the European Centre for Medium Range Weather Forecasting (ECMWF).

Despite these advances, forecast skill for drought onset, evolution and termination is often modest and highly variable across regions, seasons and lead times, which constrains the effectiveness of proactive drought early warning (Prudhomme *et al.,* 2024). As a result, many existing DEWS remain oriented toward near-real-time detection ("warn on detection") rather than anticipatory warning of emerging drought conditions and impacts. Drought events are therefore frequently declared weeks after onset, limiting opportunities for proactive risk management (Funk & Shukla*,* 2023). Uncertainty in forecasting drought persistence and termination further complicates early intervention, highlighting that effective drought management depends not only on forecast availability but also on balanced assessment of forecast skill and uncertainty within a risk-based decision-making framework (Wilhite & Vanyarkho, 2016; AghaKouchak et al., 2015).

2.3.3. Current operational DEWS worldwide

Globally, DEWS remain less developed compared to other natural hazard warning systems (such as those for floods or storms) due to the inherent complexity of drought events (UN/ISDR, 2011; Kelman & Glantz, 2014). Nonetheless, operational DEWS have been established in several regions, including North America, Europe, Africa and Australia (UNEP, 2012).

In Europe, the European Drought Observatory (EDO) provides real-time drought monitoring through maps generated using the Combined Drought Indicator (CDI), integrating the Standardised Precipitation Index (SPI), simulated soil moisture, and satellite-derived vegetation conditions (Sepulcre-Canto *et al.,* 2012). The EDO issues forecast for soil moisture anomalies up to seven days ahead and meteorological drought condition up to three months in advance (https://drought.emergency.copernicus.eu/).

In North America, the United States Drought Monitor (USDM; Svoboda *et al.,* 2002. https://droughtmonitor.unl.edu/) regularly updates drought severity assessments and changing trends by combining indices such as SPI, soil moisture observations, and satellite-derived data through experts' subjective interpretation (Hayes *et al.,* 2011). The USDM categorises drought severity into multiple classes, ranging from abnormally dry (D0) to exceptional drought (D4).

For Sub-Saharan Africa, Central America and Afghanistan, the Famine Early Warning Systems Network (FEWS-NET) issues regular drought bulletins, incorporating indicators of both drought and famine risk to enhance food security and agricultural preparedness (Senay *et al.,* 2015; Akwango *et al.,* 2017). FEWS-NET uses satellite-based indicators, including precipitation anomalies, vegetation indices, and soil moisture data, to provide current drought condition reports and monthly hazard outlooks related to food security risks (Sheffield *et al.,* 2014; Senay *et al.,* 2015).

In Australia, due to the absence of national DEWS, some state governments have established regional systems. Examples include the New South Wales (NSW) Enhanced Drought Information System (EDIS; https://www.dpi.nsw.gov.au/climate_applications/) and the Queensland Drought Monitor (https://www.longpaddock.qld.gov.au/drought/sequence/). Both systems integrate meteorological and vegetation-based metrics into a single CDI. The NSW drought indicator specifically incorporates precipitation data, soil moisture measurements, and Normalised Difference Vegetation Index (NDVI) anomalies, though current systems do not include forecast information.

In Australia, operational outlooks for pasture and agricultural drought are provided through the AussieGRASS modelling framework(https://www.longpaddock.qld.gov.au/aussiegrass/), which integrates climate, soil and vegetation information to estimate pasture growth and express conditions relative to the historical record. Built on the GRASP biophysical model (Rickert *et al.,* 2000), AussieGRASS delivers spatially explicit assessments of both current and forecast pasture conditions. Its outputs include retrospective indicators of pasture growth relative to historical percentiles, as well as seasonal outlook products such as the likelihood of exceeding median pasture growth, the most likely pasture growth tercile, and forecast curing anomalies, typically at lead times of up to three months (Carter *et al.,* 2000). Together, these probabilistic products support risk-based agricultural decision-making and operational drought assessment across Australia.

Despite these advancements, existing operational DEWS primarily focus on hazard monitoring and lack comprehensive impact-based forecasting tools. The ability to anticipate drought impacts ahead of time remains limited but is essential for proactive drought risk management (Sutanto *et al.,* 2019). For agricultural stakeholders, translating weather forecasts into actionable drought impact assessments would significantly enhance preparedness. Furthermore, impact-based forecasts could support financial institutions, such as banks, in evaluating agricultural sector vulnerability to drought. Additionally, this information could facilitate the development of innovative insurance products, particularly parametric or index-based weather insurance (Hughes, 2022).

2.3.4. Advances in seasonal drought forecasting methodologies

Drought prediction typically involves forecasting different drought phases—such as onset, severity, duration and recovery—by assessing specific thresholds of drought indices and indicators (Sharma & Panu, 2012; Wetterhall *et al.,* 2015). Current research largely focuses on forecasting meteorological and hydrological droughts, from precipitation decline to reduced streamflow, soil moisture and groundwater levels. These efforts commonly employ three primary forecasting approaches: statistical, dynamical and hybrid approaches (Hao *et al.,* 2018; AghaKouchack *et al.,* 2022).

Statistical approaches

Statistical forecasting methods fit the empirical temporally lagged relationship between drought indices and a set of potential predictors over large spatial scales (Hao *et al.,* 2018).

Various statistical methods have been applied to forecast meteorological drought, typically quantified by hydro-meteorological indices (e.g., SPI, SPEI), using predictors such as oceanic-atmospheric circulation patterns (e.g., El Niño Southern Oscillation (ENSO), the Indian Ocean Dipole (IOD)) and local climate variables (e.g., precipitation and temperature) (Dai, 2011; Yuan *et al.,* 2018). These data-driven methods do not explicitly represent physical processes and are generally most skilful at short lead times, particularly at monthly to seasonal scales, while forecast skill at annual scales is typically low and inconsistent across regions (Zhao *et al.,* 2020). Predictors, derived from hydro-meteorological observations or reanalysis data across various regions and seasons, are input into statistical modelling techniques (e.g., time series models, regression models, Markov Chain, and conditional probability models) (Hao *et al.,* 2016; Part *et al.,* 2016). Machine learning algorithms, such as Artificial Neural Networks (ANN), Random Forests (RF), and Support Vector Machines (SVM), have become widely used for modelling complex interactions between predictors and predictands (Sutanto *et al.,* 2019; Khan *et al.,* 2020).

Dynamical forecasting approaches

Dynamical (process-based) forecasting employs advanced land-surface models (LSMs) to simulate physical water cycle processes at the Earth surface-atmosphere interface (see example illustrated in Fig. 2.5). These simulations capture key variables such as precipitation, evapotranspiration, soil moisture and runoff, which collectively shape the propagation characteristics of drought (Schepen *et al.,* 2012; Fisher & Koven, 2020). Hydrological forecasts are typically driven or coupled by seasonal numerical weather forecasts from general circulation models (GCMs), such as the NCEP Climate Forecast System Version 2 (CFSv2) (Saha *et al.,* 2014) and ECMWF's seasonal forecasting system (SEAS5) (Johnson *et al.,* 2019). The soil moisture forecasts can be achieved through the North American Multi-model Ensemble (NMME) forecasts provided by NCEP's Climate Prediction Center (Becker *et al.,* 2014; Bolinger *et al.,* 2016), the LISFLOOD model developed by the ECMWF (Molteni *et al.,* 2011) and Australia's operational forecasts integrating ACCESS-S2 climate outputs into the AWRA-L hydrological model (Griffiths *et al.,* 2023), each of which deliver probabilistic seasonal outlooks for rainfall, temperature, and root-zone soil moisture conditions several months ahead (Wedd *et al.,* 2022).

Hybrid drought forecasting

Hybrid forecasting methods combine both statistical and dynamical approaches to improve model prediction accuracy (Hao *et al.,* 2018). While it is challenging to determine which method is superior, combining dynamical models with statistical models has been shown to enhance forecasting skill (Schepen and Wang, 2015; Schepen *et al.,* 2014). Common merging techniques, including regression models, Bayesian posterior distributions, and Bayesian Model Averaging (Mendoza *et al.,* 2015), are used to combine the strengths of both approaches for more reliable drought forecasting.

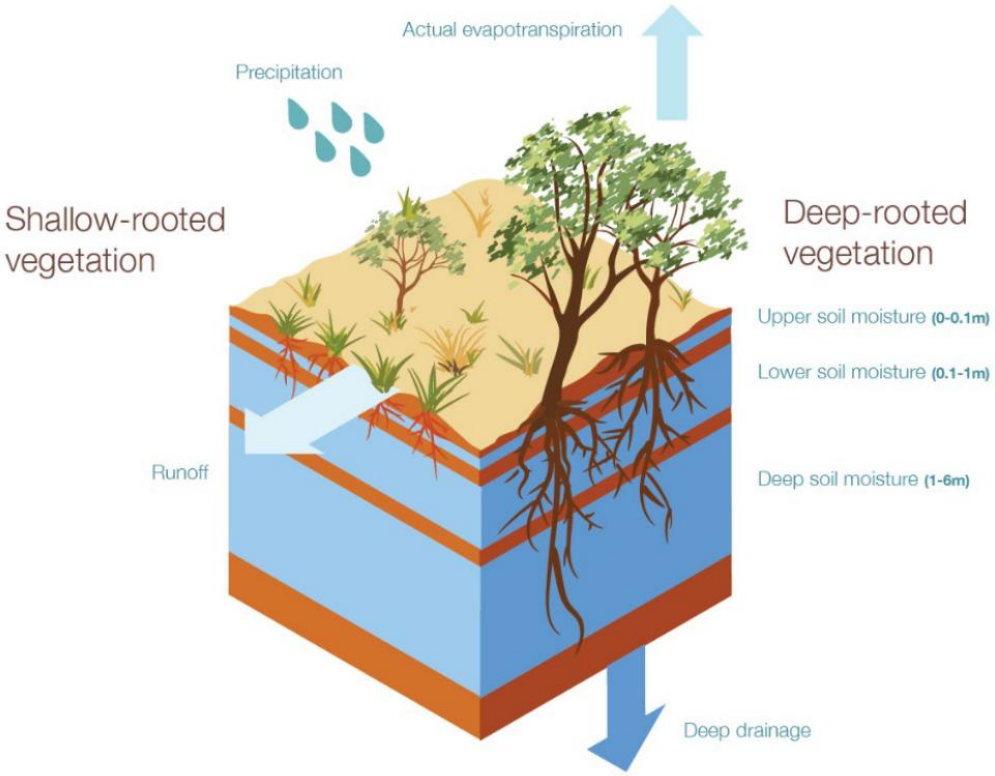


Fig. 2.5. The Australian Water Resources Assessment Landscape (AWRA-L) modelling system outputs key hydrological parameters forced by the observed meteorological variables (Van Dijk *et al.,* 2010; Viney *et al*., 2015; Frost *et al.,* 2021).

2.4. Enhancing DEWS through impact-based forecasting

2.4.1. Necessity of impact-based drought forecasting

Impact-based forecasting is crucial for improving drought early warning systems (DEWS), particularly by anticipating drought effects on vegetation and agricultural productivity. While traditional DEWS primarily focus on predicting hydro-meteorological conditions (e.g., precipitation deficits and soil moisture anomalies), impact-based forecasting explicitly connects these predictions to tangible outcomes that are directly relevant to stakeholders, enabling more effective preparedness and decision-making (Sutanto *et al.,* 2019).

Despite their recognised importance, operational DEWS currently lack robust, probabilistic forecasts of drought impacts. Existing systems often fail to transform drought signals from meteorological forecasts into specific, actionable information tailored to end-user needs. Although the terms "drought forecasting" and "drought early warning" are sometimes used interchangeably, they differ in scope. Drought forecasting typically refers to predicting the likelihood, onset and severity of drought conditions, whereas drought early warning extends this by disseminating timely, practical information to stakeholders, facilitating proactive risk mitigation measures. Thus, incorporating direct indicators of agricultural and ecological drought impacts—such as vegetation stress or reduced agricultural yields—into DEWS can significantly enhance their effectiveness compared to systems focused exclusively on hydro-meteorological indices.

A primary goal of impact-based DEWS is to supply accurate information on drought impacts to governments, agricultural stakeholders and financial institutions. For agriculture, impact-based forecasting commonly integrates climatic forecasts with dynamic agronomic models to predict biophysical outcomes like crop yields and pasture growth. For example, in Australia, the Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES) is developing national indicators that combine climate data with agricultural modelling. These indicators translate forecasts into meaningful agricultural outcomes, such as anticipated crop yields, pasture conditions, and broader farm economic impacts, supporting more informed decision-making within the agricultural and financial sectors (Hughes *et al.,* 2022).

2.4.2. Advances in forecasting vegetation impacts for DEWS

Compared to established methods for forecasting meteorological and hydrological drought, the forecasting of drought impacts—particularly those affecting vegetation—is relatively underdeveloped (Bachmair *et al.,* 2016). While several studies have attempted to predict drought signals using various drought indices, few have explicitly addressed their direct translation into ecological or socio-economic impacts (Bachmair *et al.,* 2016; Sutanto *et al.,* 2019).

Among existing efforts, most studies primarily focus on crop yield as the main metric to quantify agricultural drought impacts. Such yield-based forecasting typically combines meteorological forecasts (e.g., three-month SPI) or observed vegetation indices (e.g., NDVI anomalies) with agronomic models to provide insights into expected agricultural productivity (Peng *et al.,* 2018, 2020; Krishnamurthy *et al.,* 2020). Beyond yield forecasting, studies on predicting socio-economic drought impacts remain limited and geographically restricted, typically relying on drought-impact databases such as the U.S. Drought Impact Reporter

(DIR) and the European Drought Impact Inventory (EDII). These text-based impact inventories compile records of drought-induced impacts on sectors including domestic water supply, agriculture, hydropower and water quality (Bachmair, Stahl, *et al.,* 2016; González Tánago *et al.,* 2016; Stagge *et al.,* 2015). For instance, Sutanto et al. (2019) used logistic regression and Random Forest methods based on EDII data to predict drought-related impacts across Europe. However, such inventories are not universally available, thereby restricting broader applicability.

Predicting drought impacts on vegetation at large scales faces multiple challenges, notably due to the complexities of vegetation responses to varying water stress across different climate zones (Sheffield & Wood, 2011). Remote sensing data, particularly NDVI and related vegetation indices, are frequently employed as proxies for vegetation conditions and have been commonly forecast using statistical models (Asoka & Mishra, 2015; Funk & Brown, 2006; Tadesse *et al.,* 2014). Although previous approaches such as linear autoregression and Gaussian process (Barrett *et al.,* 2020), linear regression (Zambrano *et al.,* 2018) and Bayesian autoregression (Salakpi *et al.,* 2022) have attempted to forecast vegetation condition using single-variable models based solely on individual vegetation indices with short lead times (1–2 months), their temporal coverage remains insufficient for operational drought early warning and effective preparedness.

Statistically-based forecasts typically utilise regression and machine learning algorithms, employing hydro-meteorological predictors such as precipitation and temperature or large-scale climate indices (e.g., ENSO) to estimate vegetation responses to drought (Hao *et al.,* 2018). Dynamic forecasting approaches further integrate vegetation indicators (e.g., NDVI) and hydro-meteorological variables through data assimilation within land-surface models, enabling near-real-time monitoring and short-term prediction of agricultural drought conditions (Bolten & Crow, 2012; Tian *et al.,* 2019). Nevertheless, most studies traditionally focus on standard meteorological indices (e.g., SPI, SPEI) rather than soil moisture anomalies, despite the latter's direct relevance to vegetation growth, particularly in water-limited dryland ecosystems (AghaKouchak *et al.,* 2015).

Recent Australian studies have demonstrated the practical value of explicitly forecasting soil moisture for drought and vegetation-related applications. For example, Western *et al.* (2018) evaluated seasonal forecasts of plant available soil water across dryland cropping systems in southeastern Australia and found skill exceeding climatology at lead times of up to 3–5 months. At the continental scale, Vogel *et al.* (2021) showed that a national seasonal ensemble forecasting system based on AWRA-L exhibited useful skill for soil moisture and evapotranspiration forecasts at one- to three-month lead times across key agricultural regions, with particularly high skill for low soil moisture extremes. These findings highlight the potential of soil moisture–based forecasting to support vegetation impact assessment in Australian dryland systems, where soil moisture exerts a primary control on vegetation dynamics.

Recent studies have begun addressing this gap by explicitly linking soil moisture predictions to vegetation conditions such as NDVI, LAI and GPP. Methods employed include regression analysis, conditional probability modelling, and machine learning algorithms to establish robust predictive relationships (Asoka & Mishra, 2015; Sawada & Koike, 2016; Tian *et al.,* 2019). Notably, soil moisture anomalies derived from in situ measurements (Gu *et al.,* 2008; Wang *et al.,* 2007), satellite remote sensing products (Chen *et al.,* 2014; Nicolai-Shaw *et al.,* 2017), and land surface model simulations (Ahmed *et al.,* 2017; Crow *et al.,* 2012) often exhibit stronger predictive relationships with vegetation stress than meteorological indices alone, especially in dryland regions like Australia (Bachmair *et al.,* 2016a; Chatterjee *et al.,* 2022). Several studies have quantified predictive potential using temporal correlation measures, such as Pearson's or Spearman's rank correlations and Anomaly Correlation (AC), between satellite-derived vegetation indices and hydro-meteorological variables (Sehgal *et al.,* 2021a). However, correlation-based measures alone may not adequately indicate forecast skill, especially given the rarity and severity of extreme drought events, which fall in the distribution tails of both vegetation indices and soil moisture data (Wilhite & Glantz, 1985). Consequently, predictive skill derived from historical correlation often deteriorates when applied to forecasting statistically rare drought events, presenting persistent methodological challenges (Brust *et al.,* 2021; Hao *et al.,* 2017; Sutanto *et al.,* 2020).

Recent Australian research has demonstrated the value of explicitly forecasting soil moisture for drought and vegetation impact applications. For example, seasonal ensemble forecasting systems driven by climate model outputs and land surface models (e.g., AWRA-L) show useful skill for soil moisture at lead times of one to three months across key agricultural regions, with particularly strong performance for low soil moisture extremes (Western et al., 2018; Vogel et al., 2021).

2.4.3 Probabilistic forecasting and decision-making under uncertainty

Given the inherently chaotic nature of the climate system and limitations of current predictive capabilities, stakeholders often rely on probabilistic forecasts, which explicitly quantify prediction uncertainty, to support decision making (Demargne *et al.,* 2014; Hao *et al.,* 2017). Consequently, probabilistic drought forecasts that clearly present associated uncertainties have become critical components of effective drought early warning systems (DEWS).

Uncertainty in drought forecasts is generally quantified through probabilistic methods, such as the derivation of probability density functions (PDFs) of forecast variables (e.g., precipitation or drought indices) by combining ensemble forecasts from multiple models (Stockdale *et al.,* 2010). In practice, the mean, variance, or ensemble spread are common metrics used to represent forecast uncertainty (Yuan & Wood, 2012). Compared to deterministic forecasts, probabilistic forecasts often provide greater value, offering clearer communication of uncertainty and supporting better-informed decisions (Hao *et al.,* 2018). Recent advancements in dynamical forecasting enable the generation of probabilistic drought predictions using multi-model ensemble systems, typically focused on key hydro-meteorological variables such as precipitation or soil moisture.

A critical aspect of drought early warning is the lead time of forecasts, which directly influences decision-making effectiveness. However, longer lead times generally come with increased forecast uncertainty, creating an inherent trade-off between timely warnings and reliability (Grasso, 2006). Existing studies suggest that drought hazards can exhibit useful but variable forecast skill at lead times of approximately two to three months, depending on region, variable and drought metric considered (Sutanto *et al.,* 2019a).

Soil moisture is particularly important for defining and monitoring agricultural drought impacts due to its direct influence on plant growth and productivity, especially in water-limited ecosystems (Wilhite, 2005). Nevertheless, the use of soil moisture in operational drought forecasts faces significant challenges. Soil moisture exhibits higher spatial variability than precipitation, and comprehensive, large-scale historical observational datasets are scarce (Dorigo *et al.,* 2017). These constraints hinder the development and operationalisation of reliable soil moisture-based drought forecasts, consequently limiting the effectiveness of early warnings for agricultural drought management.

To address these constraints, this thesis contributes a soil moisture–driven prediction framework that integrates both deterministic and probabilistic elements by leveraging state-of-the-art root-zone soil moisture simulations and ensemble forecasts, with the aim of enhancing early warning capabilities for drought-induced vegetation stress in operational contexts.

Chapter 3

Categorical prediction of vegetation deterioration events using large-scale root-zone soil moisture anomalies over southeastern Australia

Abstract

Predicting vegetation response to drought stress remains challenging due to complex interactions between soil moisture deficits and plant health, compounded by varying response times and magnitudes across diverse ecosystems. This study develops a probabilistic framework for determining soil moisture trigger thresholds associated with multi-categorical vegetation deterioration risks in southeastern Australia. The framework integrates time-lagged correlation analysis with copula-based joint probability modelling to establish quantitative relationships between root-zone soil moisture anomalies and vegetation condition. Analysis of satellite-derived vegetation indices and Australian Water Resources Assessment – Landscape model (AWRA-L) simulated soil moisture (2001-2022) reveals strong temporal coupling between vegetation dynamics and root-zone moisture conditions, with response times typically ranging from 8 to 40 days. The copula-based analysis successfully captures soil moisture-vegetation dependencies, enabling probabilistic estimation of vegetation deterioration under varying drought scenarios. Even under severe soil moisture deficits ($\leqslant$ -2.5 SD), the region-wide probability of experiencing mild vegetation deterioration averages only 61.2%, highlighting the complex nature of drought-vegetation interactions. The predictive capability of this framework is evaluated based on its performance using multiple conditional probability criteria (40-70%) for triggering vegetation deterioration alarms under specific soil moisture anomaly conditions. Results demonstrate strong skill in forecasting no-risk to moderate vegetation stress conditions (hit rates: 0.87 and 0.33 respectively) but diminished performance for severe categories. Agricultural regions, particularly dryland cropping areas, exhibit extended response times exceeding one month, while natural vegetation systems show more rapid responses to moisture stress. This research advances impact-based drought early warning by providing a systematic approach for quantifying vegetation deterioration risks. The methodology offers valuable insights for agricultural and environmental management, though challenges remain in predicting extreme vegetation stress events.

3.1. Introduction

3.1.1. Motivation and problem context

As emphasised in Chapter 2, drought is one of the costliest natural hazards, exerting significant impacts on both human societies and ecosystems (Mishra and Singh, 2010). When drought conditions propagate into significant soil moisture deficits in the effective root zone (0-1m, as represented by large-scale land surface models), they can severely impair vegetation functioning by reducing photosynthetic activity, suppressing growth, and potentially inducing plant mortality (Dai, 2011; Farooq *et al.,* 2009; Seleiman *et al.,* 2021). These effects may cascade into broader consequences. For example, crop yields and harvest quality are highly sensitive and may be reduced very early in a stress episode, often before the full stress is expressed in the overall plant (Chawla et al., 2020; Rajsekhar et al., 2015). Other impacts include deteriorating pasture conditions for grazing (Yang *et al.,* 2023), decreased terrestrial carbon stocks, reduced ecosystem services as ecological drought (Crausbay *et al.,* 2017), and increased bush/wild fire risk (Jolly *et al.,* 2015). Vegetation degradation is known to be closely linked to soil moisture stress, especially in dryland ecosystems (Zhang *et al.,* 2023).

Predicting drought impacts on vegetation remains a major challenge due to two key issues, as also introduced in Chapter 2. First, vegetation exhibits non-linear and delayed responses to hydro-meteorological variability, with sensitivity and vulnerability varying significantly across different ecosystems and climatic regimes (Vicente-Serrano *et al.,* 2013; Sun *et al.,* 2021). Soil moisture deficits do not always result in vegetation deterioration, particularly when plant drought resistance mechanisms buffer stress until certain thresholds are surpassed (Frank *et al.,* 2015; X. Li *et al.,* 2023; Liu *et al.,* 2019). Second, the characteristics of drought events (e.g., timing, intensity and duration) critically influence vegetation responses (Frank *et al.,* 2015). For example, droughts occurring near peak growth periods can disproportionately reduce ecosystem productivity (Jiao *et al.,* 2022). These complexities in vegetation response to drought have motivated a range of statistical approaches aimed at characterising soil moisture–vegetation relationships. The following subsection reviews key methodologies developed to analyse these dynamics.

3.1.2. Existing approaches to soil moisture-vegetation analysis

As discussed in Chapter 2, drought impact assessment has increasingly relied on remote sensing indicators and hydro-meteorological variables to monitor vegetation responses at large spatial scales. Building on this foundation, the present chapter focuses on statistical approaches developed to characterise the relationship between soil moisture deficits and vegetation deterioration, particularly those suited for identifying drought propagation dynamics and impact thresholds. Broadly, three methodological approaches have been applied in recent studies: (1) correlation analysis, (2) joint probability modelling, and (3) coincidence analysis.

Time-lagged correlation analysis estimates the temporal response of vegetation to drought stress by calculating correlation coefficients (e.g., Pearson, Spearman's rank) between vegetation dynamics proxies and drought indices at various lags (Vicente-Serrano *et al.,* 2013; Zhang *et al.,* 2022b). This method captures both the overall correlation strength and timescale of vegetation response to hydro-meteorological drought condition but is limited in its ability to detect dependencies in extreme conditions (Schwalm *et al.,* 2017, Yu *et al.,* 2017). It often fails to capture tail dependence, which is critical for understanding severe drought impacts on vegetation (Zscheischler *et al.,* 2018). In this study, time-lagged Spearman correlation analysis is used as a preliminary step to quantify the temporal alignment between root-zone soil moisture anomalies and vegetation deterioration across southeastern Australia.

Thus, a logical alternative is to focus on the lower tail of the distribution (herein referred to as the ‘dry tail’) of hydro-meteorological and vegetation variables, where drought-related anomalies typically manifest. Researchers have increasingly adopted copula-based joint probability modelling, which allows flexible modelling of non-linear dependence and extreme event co-occurrence by calculating joint exceedance (or non-exceedance) probabilities (Hao *et al.,* 2016, 2023). For instance, this approach constructs joint distributions between soil moisture and vegetation variables (e.g., NDVI anomalies) and can answer conditional probability questions, such as: “Given a root-zone soil moisture anomaly below –1.0 SD, what is the probability of vegetation deterioration exceeding –1.5 SD?” (Sklar, 1959; Ribeiro *et al.,* 2019). Two main directions have emerged in this field: (1) assessing vegetation vulnerability by analysing tail dependence structures (Fang *et al.,* 2019; Xu *et al.,* 2023), and (2) identifying critical drought thresholds linked to specific levels of vegetation loss (Faiz *et al.,* 2022; He *et al.,* 2021). This study adopts bivariate copula-based modelling as a core component of its prediction framework, using it to quantify conditional probabilities and determine critical soil moisture thresholds for multiple categories of vegetation deterioration.

A third approach, coincidence analysis, evaluates the co-occurrence rate between extreme soil moisture deficits and vegetation anomalies. Typically, percentile-based thresholds (e.g., NDVI anomalies <10th percentile, soil moisture anomalies < 20th percentile) are used to define events, and their co-occurrence rate is computed across different timescales, from sub-monthly (Tang *et al.,* 2024) to annual periods (Sun *et al.,* 2023a; Yang *et al.,* 2023a; Zhang *et al.,* 2023). This method has been applied to both regional (Sun *et al.,* 2023) and global (X. Li *et al.,* 2023; Tang *et al.,* 2024) assessments of vegetation sensitivity to soil moisture stress. While coincidence analysis provides useful insights into the empirical relationship between soil moisture and vegetation anomalies, it is not applied in this study. Instead, I focus on conditional probability-based methods that offer stronger predictive capabilities suited for early warning applications.

3.1.3. Limitations and research gap

Although these approaches have advanced our understanding of drought propagation to vegetation systems, significant gaps remain, particularly in predicting severe and low-frequency vegetation deterioration events that are crucial for effective impact-based drought early warning systems. While previous studies have explored both overall patterns and tail dependence, their predictive accuracy for the most severe events remains uncertain. The identified thresholds in prior studies typically range from 40^{th} to 10^{th} percentiles (or, where a Gaussian distribution is assumed, z-scores from -0.25 SD to -1.28 SD), leaving the predictability of more extreme deterioration (e.g., z-scores < -1.5 SD or < -2.0 SD) largely unaddressed.

Moreover, much of the existing research treats “agricultural drought” as a composite category, conflating soil moisture deficits with their vegetation impacts. However, vegetation responses are not deterministic: soil moisture deficits do not necessarily result in vegetation deterioration of a corresponding severity (Sun *et al.,* 2023). The likelihood and extent of vegetation impact depend on multiple factors, including plant functional traits, ecosystem type, timing of the drought event, and cumulative exposure (Anderegg *et al.,* 2020). This underscores the inherently probabilistic nature of soil moisture–vegetation relationships, and highlights the need to model these components separately and explicitly. In addition, most prior studies have focused on agricultural systems due to their economic significance. Fewer have assessed drought impacts on native vegetation, despite its importance for ecosystem function, biodiversity and land management (Retallack *et al.,* 2023).

3.2. Objective and contribution

To address the limitations identified above, this study analyses the probabilistic relationship between soil moisture drought and vegetation deterioration through capturing not only whether an impact occurs, but also how likely and severe it may be under varying drought conditions. I develop pixel-wise soil moisture anomaly (SMA) thresholds for predicting vegetation deterioration (VD) risks using copula-based joint probability modelling. I adopt the Normalized Difference Vegetation Index (NDVI) as the primary indicator due to its strong association with leaf biomass and its suitability for large-scale remote sensing. This study advances existing approaches by constructing a predictive skill evaluation framework that leverages bivariate copula models to estimate VD risks from SMA patterns at the pixel scale across southeastern Australia. Specifically, I aim to:

1. Characterise the temporal relationship and tail-dependence structure between vegetation dynamics and root-zone soil moisture variability using time-lagged correlation and bivariate copula modelling at the pixel level.
2. Estimate joint occurrence probabilities and determine localised, pixel-specific critical soil moisture thresholds for triggering multiple severity levels of vegetation deterioration.
3. Evaluate the predictive skill of these trigger thresholds, with a particular focus on extreme and low-frequency events to enhance early warning capabilities.

By targeting both agricultural and native vegetation systems, this research contributes to the development of robust early warning tools that support improved drought preparedness and impact mitigation.

3.3. Study area and datasets

The study area encompasses southeastern Australia (33°S–40°S, 140°E–150°E), a region characterised by recurring drought events in recent decades. The analysis period focuses on austral late autumn to middle spring (May to October), coinciding with the primary precipitation season and peak vegetation growth phase in the wheat-sheep belt of southeastern Australia (Chenu *et al.,* 2013; Donohue *et al.,* 2018; Feng *et al., 2020*). This temporal window was selected to optimise the investigation of soil moisture drought effects on vegetation dynamics.

The study utilises root-zone (0-1.0m) soil moisture data derived from the Australian Water Resources Assessment Landscape model (AWRA-L, version 7) operational historical simulation for the period 2001-2022. These water balance outputs were obtained from the Australian Bureau of Meteorology's Australian Water Outlook (https://thredds.nci.org.au/thredds/catalog/catalogs/iu04/australian-water-outlook/historical/historical.html). The daily soil moisture outputs underwent temporal aggregation using an 8-day surrounding window mean to align with satellite-derived vegetation index temporal resolution. Initial statistical analysis revealed significant positive skewness in the root-zone soil moisture estimates, particularly pronounced in central regions dominated by dryland cropping and pasture. To address this non-normality, a square root transformation was applied at the pixel level. Subsequent Shapiro-Wilk testing confirmed improved normality in the transformed dataset hereafter referred to as RZSM.

Vegetation conditions were quantified using 8-day Normalized Difference Vegetation Index (NDVI) time series retrieved from the Moderate Resolution Imaging Spectroradiometer (MODIS) nadir BRDF-adjusted reflectance product (MCD43A4 Collection 6, https://lpdaac.usgs.gov/products/mcd43a4v006/). The analysis incorporated only highest-quality retrievals of calculated bands 1 and 2. The NDVI dataset underwent spatial resampling to 0.05°×0.05° resolution using mean aggregation, ensuring spatial correspondence with soil moisture data coordinates. The processing workflow included mosaicking of sinusoidal tile NDVI data for the study area and reprojecting to a regular geographic grid through bilinear interpolation. Temporal gaps in the NDVI series were addressed through linear interpolation, with gap-filled data constituting only 4% of the complete time series across the study area and analysis period.

The analysis incorporated land-use data from the Australian Collaborative Land Use and Management Program (ACLUMP version 8, https://www.agriculture.gov.au/abares/aclump), originally mapped at 50m resolution. The dataset was reprojected to the 0.05° (~5km) grid of the geographic coordinates of the AWRA-L, and each grid cell was assigned using a majority-rule aggregation, whereby the most frequent ACLUMP land-use class within the grid cell was taken as the representative type. The primary ACLUMP hierarchy was used, encompassing major categories such as grazing modified pastures, dryland cropping, grazing native vegetation, nature conservation, production native forests, plantation forests, irrigated pasture and irrigated cropping (Fig. 3.1a). ACLUMP v8 provides a static snapshot for 2016; thus, temporal variations in land use were not incorporated. To minimise sub-grid heterogeneity in vegetation characteristics, the analysis was restricted to pixels in which a single land use type occupied more than 50% of the cell area, encompassing 85.2% of all vegetation-dominated grid cells. This land-use stratification enable evaluation of predictive capability across distinct vegetation regimes and facilitates the development of targeted drought management and preparedness strategies.

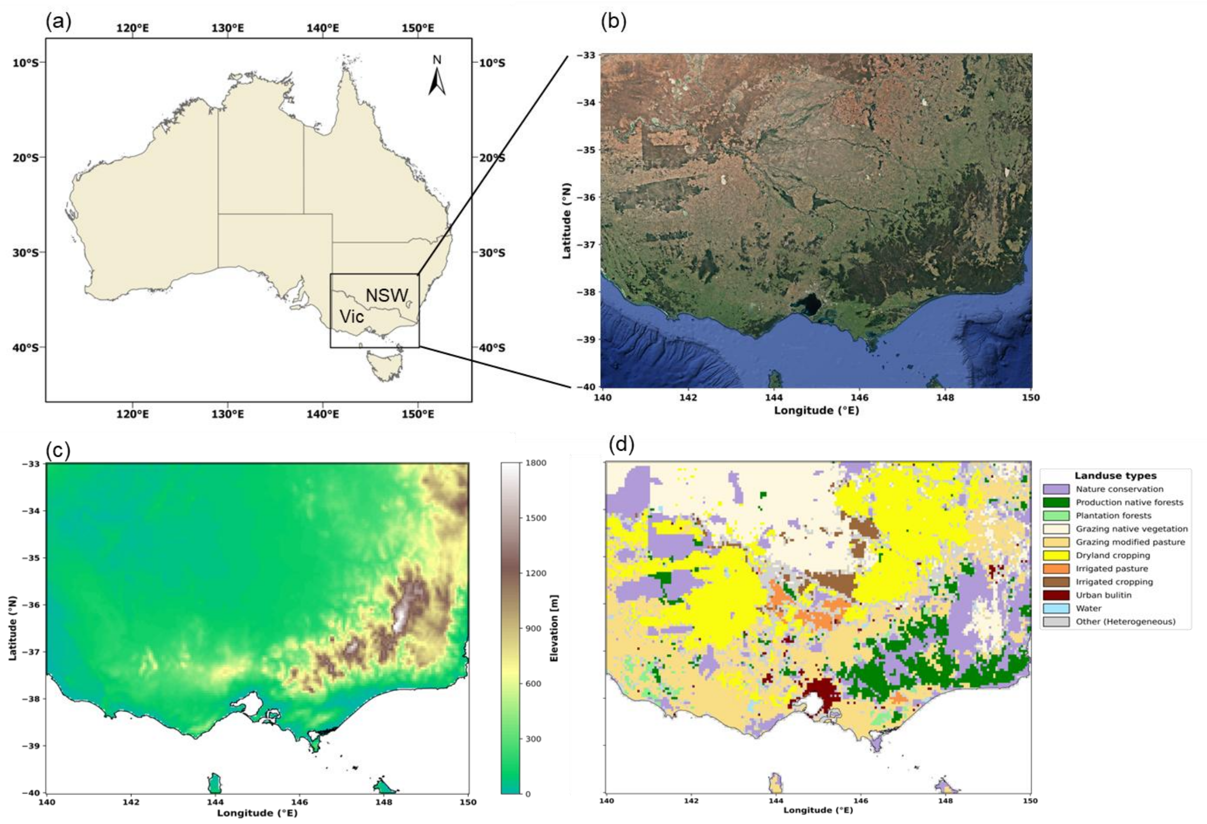


Fig. 3.1. Overview of the study area in southeastern Australia. (a) Location of the study region, encompassing parts of New South Wales (NSW), Victoria (Vic) and Australian Capital Territory (ACT). (b) True-colour satellite image of the region (Google Earth). (c) Elevation map of the study area derived from a digital elevation model. (d) Land use and vegetation types across the region, classified into major land use categories.

3.4. Methodology

This section outlines the methodological framework developed to predict vegetation deterioration (VD) severity based on standardised soil moisture anomalies (SMA). The complete workflow is illustrated in Fig. 3.2, with corresponding subsections detailing each analytical step. The approach is structured into five major stages, each designed to address a specific component of the forecasting framework.

In Step 1 (Section 3.4.1), I characterise anomalies in vegetation and root-zone soil moisture conditions using standardised NDVI and RZSM time series. Step 2 (Section 3.4.2.1) identifies the temporal response lag of vegetation to soil moisture deficits by estimating the time-shifted correlation between SMA and VIA series. Building on this, Step 3 (Sections 3.4.2.2–3.4.2.3) constructs pixel-wise bivariate dependence models using copula functions. Four candidate marginal distributions and four copula families are evaluated, and the best-fitting copula for each pixel is selected using the Kolmogorov–Smirnov (K–S) test, the Cramér–von Mises (CvM) statistic, and the Akaike Information Criterion (AIC).

In Step 4 (Section 3.4.2.4), I derive conditional cumulative probabilities of vegetation deterioration from the fitted copula models, which are used to identify SMA trigger thresholds for different severity levels of VD. These thresholds are determined under varying exceedance probability criteria (40%, 50%, 60%, and 70%), producing four prediction sets: Copula-cp40, cp50, cp60, and cp70. Each set shares the same underlying pixel-specific copula function, differing only in the decision threshold applied. Finally, Step 5 (Sections 3.4.3–3.4.4) evaluates the predictive performance of the framework by comparing copula-based forecasts to a naïve baseline (persistence prediction), using historical SMA time series across the study region. Predictive performance skill is assessed using multiple metrics, including RMSE, hit rate, false alarm rate, and frequency distribution.

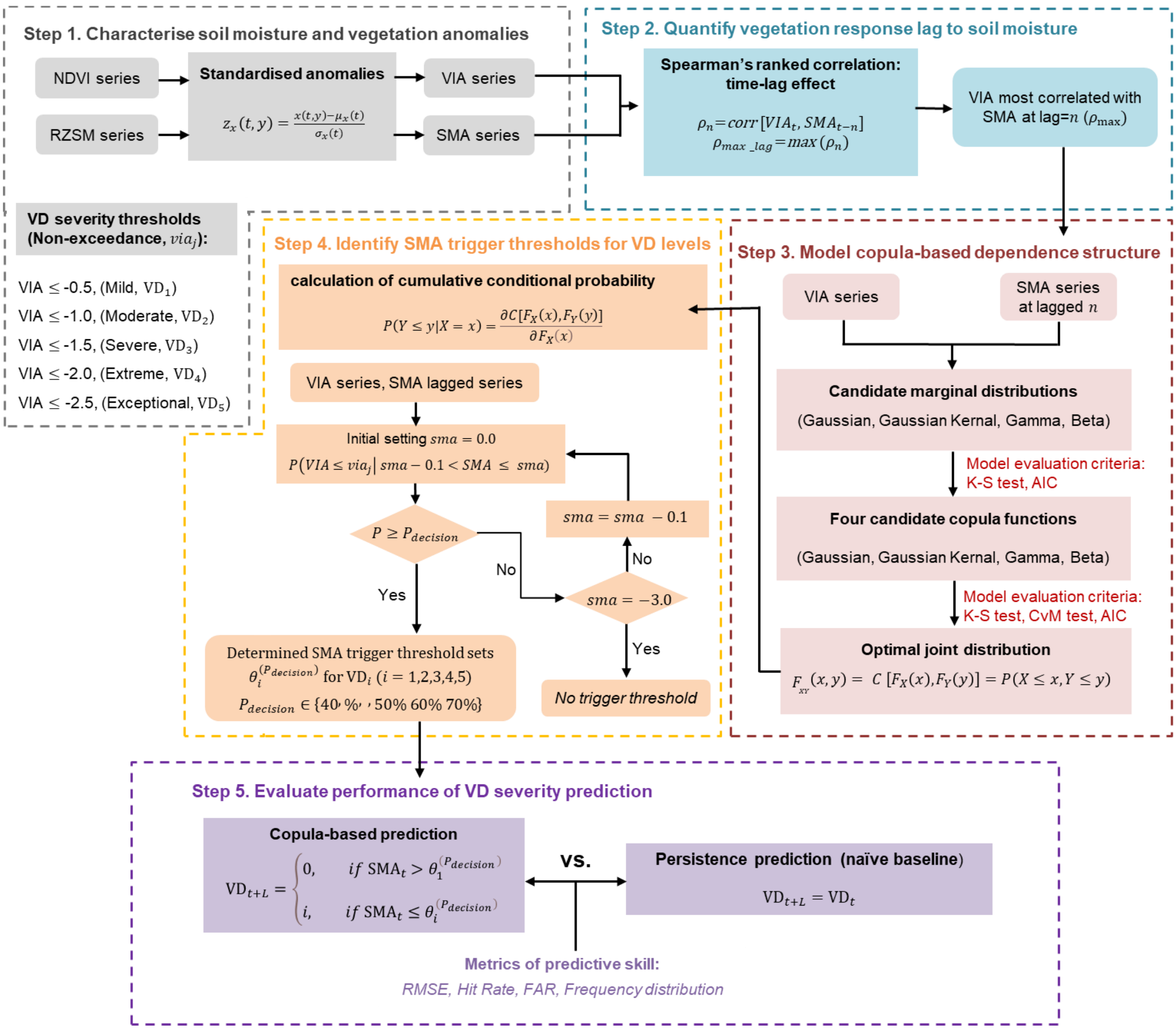


Fig. 3.2. Pixel-wise workflow for deriving SMA thresholds associated with varying vegetation deterioration (VD) severity levels, using copula-based conditional probability modelling.

3.4.1. Standardised vegetation and soil moisture anomalies

I derived standardised anomalies (z-scores) of these variables for each grid using the following equation:

$$z_x(t, y) = \frac{x(t,y) - \mu_x(t)}{\sigma_x(t)} \tag{3.1}$$

where $x(t, y)$ represents the value of a variable for day t of the year y in the 2001-2022 period, and $\mu_x(t)$ and $\sigma_x(t)$ denote the multi-year mean and standard deviation of the variable x at day *t* over the study period, respectively. The calculation of $x(t, y) - \mu_x(t)$ yields the anomaly for day *t* of year *y*, which is then standardised to $z_x(t, y)$.

This standardisation procedure de-seasonalises the original variables and transforms them into locally comparable z-scores. I designate the z-score of RZSM anomaly as Soil Moisture Anomaly (SMA) and the z-score of NDVI anomaly as Vegetation Index Anomaly (VIA). The z-score approach was selected over percentile-based methods due to its superior capacity to preserve distributional properties and detect extreme events more effectively, while also allowing consistent comparisons across different drought indices over spatial and temporal scales (Hao & AghaKouchak, 2013; Farahmand & AghaKouchak, 2015). In addition, z-scores remain sensitive to subtle shifts in vegetation and soil moisture responses, which is particularly important for capturing the progressive onset and evolution of drought impacts on vegetation.

The severity classification scheme for both SMA and VIA adopts the following non-exceedance threshold criteria: below-normal (≤ 0 SD), mild (≤ −0.5 SD), moderate (≤ −1.0 SD), severe (≤ −1.5 SD), extreme (≤ −2.0 SD), and exceptional (≤ −2.5 SD). Unlike conventional drought index severity classification systems defined by mutually exclusive upper and lower bounds (e.g., −1.5 < SPEI ≤ −1.0), this framework independently triggers each severity level when the anomaly value drops below a given threshold. This structure is more suitable for early warning applications (WMO & GWP, 2016; Bachmair *et al.,* 2016), where the primary goal is to detect whether a critical impact threshold has been breached. Moreover, single-sided thresholding supports the use of unidirectional exceedance probabilities in copula-based modelling, particularly for rare but severe events where traditional two-tailed classifications are less informative.

3.4.2. Copula-based framework for determining soil moisture trigger thresholds associated with vegetation deterioration

This section introduces a probabilistic framework for quantifying the relationship between soil moisture deficits and vegetation deterioration through copula-based analysis. The framework leverages the mathematical properties of copula functions to construct joint probability distributions between SMA and VIA, enabling the identification of critical soil moisture trigger thresholds associated with varying severity levels of vegetation stress.

The theoretical foundation of my approach is grounded in Sklar's Theorem (1960), a fundamental principle in statistical theory that demonstrates how multivariate distributions can be separated into two distinct components: their univariate marginal distributions and a copula function (Durante & Sempi, 2015). The copula function specifically characterises how these variables interact and depend on each other. Notably, copulas excel at capturing non-linear relationships between variables whilst preserving the unique statistical properties of each individual variable's distribution (Mishra & Datta-Gupta, 2017; Chen & Guo, 2019). This capability is believed valuable when modelling the complex interactions between soil moisture deficits and vegetation response patterns (Xiang *et al.,* 2023).

3.4.2.1. Temporal response analysis of soil moisture-vegetation relationships

Correlation analysis is the foundation for constructing a copula function as it establishes the statistical foundation for subsequent copula modelling and determines the optimal temporal scale at which soil moisture anomalies influence vegetation condition. It is reported that the relationship between vegetation dynamics and root-zone soil moisture variabilities exhibits strong non-linear characteristics, necessitating the use of rank correlation analysis methods (e.g., Kumar *et al.,* 2016; Tian *et al.,* 2019).

The Spearman's rank correlation coefficient (ρ) was calculated between the VIA and SMA time series during the growing season, incorporating various time lags. The temporal analysis spans –64 to +64 days, corresponding to eight 8-day intervals before and after the target vegetation index date, thus capturing both antecedent (leading) and delayed (lagging) soil moisture effects. Because the analysis is centred on the VIA observations within the May–October window, SMA values were obtained by shifting the SMA time series forward or backward by the corresponding lag. This approach ensured full lag coverage for all VIA dates without discarding edge observations. The relationship is expressed as:

$$\rho_n = corr\ [VIA_t, SMA_{t-n}],\ -8 \le n \le +8 \tag{3.2}$$

$$\rho_{max\_lag} = max(\rho_n),\ -8 \leq n \leq +8 \tag{3.3}$$

where *n* represents the temporal offset in $n \cdot 8$-day periods, SMA_{t-n} denotes the soil moisture anomaly *n* temporal steps before time *t*, and ρ_n represents the corresponding correlation coefficient between VIA and lagged SMA. Only pixels where the ρ_{max_lag} was found to be statistically significant ($p < 0.01$) were carried forward for the bivariate copula modelling. This criterion ensures the robustness of the derived dependency structure by filtering out regions where the vegetation dynamics are weakly or randomly coupled with root-zone soil moisture.

For each grid cell, the maximum correlation coefficient (ρ_{max_lag}) and its corresponding lag period were identified. The ρ_{max_lag} value indicates the strongest temporal relationship between soil moisture and vegetation response at that location. This optimal lag time was subsequently incorporated into the joint distribution modelling, ensuring that the temporal dynamics of vegetation response to soil moisture deficits were appropriately captured in the copula-based analysis.

3.4.2.2. Copula-based joint probability distribution

The copula framework provides a robust statistical foundation for modelling complex dependence structures between random variables by decomposing joint probability distributions into their marginal components (Varol *et al.,* 2023; Wang *et al.,* 2022). This framework offers flexibility in connecting diverse marginal distributions via specified copula functions (Nelsen, 2006; Hao *et al.,* 2017). Within hydro-meteorological research, copula-based joint distributions have proven effective in quantifying bivariate and multivariate relationships among drought-related variables across spatiotemporal scales, notably in capturing tail dependence structures essential for extreme events (Zscheischler & Seneviratne, 2017; Hao *et al.,* 2023).

The framework capability in constructing conditional probability facilitates rigorous investigation of environmental variable interactions, thereby enabling predictand-predictor relationships to be evaluated. This approach proves especially valuable in my analysis of soil moisture-vegetation relationships, where I employed copula functions (Eq. 3.4) to establish joint distributions between SMA and VIA series at empirically determined optimal time lags. For my bivariate copula modelling of SMA-VIA relationships, the joint cumulative distribution function is expressed as:

$$F_{XY}(x, y) = C\ [F_X(x), F_Y(y)] = P(X \leq x, Y \leq y) \tag{3.4}$$

where $F_{XY}(x, y)$ represents the joint cumulative distribution function (CDF), *C*[ ] donates the copula function, and $F_X(x)$ and $F_Y(y)$ are the marginal CDFs of SMA and VIA, respectively. The marginal CDFs, $F_X(x)$ and $F_Y(y)$, are uniformly distributed on the interval [0, 1], whilst the copula function *C*[ ] captures the underlying dependency structure between the variables. Finally, this expression is equivalent to $P(X \leq x, Y \leq y)$, which quantifies the joint probability that X does not exceed x and Y does not exceed y simultaneously.

In the implementation phase, I evaluated four candidate marginal distributions (Gaussian, Gaussian kernel, Gamma, and Beta) for fitting the time series of SMA and VIA. The joint dependence structure between SMA and VIA was modelled using four established copula families: Clayton, Gumbel, Frank, and Gaussian. These represent two distinct copula types, elliptical (Gaussian) and Archimedean (Clayton, Gumbel, Frank), each characterised by different tail dependence structures (Guo *et al.,* 2021; Sadegh *et al.,* 2017). The Gaussian and Frank copulas are particularly suited for capturing both positive and negative dependencies across the bivariate distribution.

To account for the spatial heterogeneity in SMA–VIA relationships, all distribution fitting and copula modelling procedures were implemented independently at each pixel. This pixel-wise fitting strategy enabled the capture of localised distributional characteristics and joint dependency structures across the study region. The selection of optimal marginal distributions and copula functions followed a rigorous statistical procedure, as illustrated in the workflow diagram (Fig. 3.2). For each pixel, marginal distributions were evaluated using the Kolmogorov–Smirnov (K–S) test and the Akaike Information Criterion (AIC), balancing empirical fit with model simplicity (Fang *et al.,* 2019). Copula functions were selected based on AIC and two goodness-of-fit metrics, the K–S test and the Cramér–von Mises (CvM) statistic, to ensure both penalised likelihood performance and consistency with the observed joint behaviour. Further details on the candidate distributions and copula families are provided in Tables S3.1 and S3.2 in the Supporting Information.

Fig. 3.3 illustrates the fundamental principles of copula theory by demonstrating how bivariate marginal distributions are transformed into their uniform counterparts over the interval [0,1], while preserving the inherent dependence structure between the variables. The original bivariate probability space (Fig. 3.3a) is divided into four quadrants (I-IV), representing the joint probability regions: $P(X > x, Y > y)$, $P(X \leq x, Y > y)$, $P(X \leq x, Y \leq y)$, and $P(X > x, Y \leq y)$. These four quadrant probabilities sum to unity, with conditional probabilities derivable through integral computation (Yue and Rasmussen, 2002). For instance, the shaded area in quadrant III corresponds to the joint probability of a drought event where soil moisture anomaly (SMA) ≤ –0.9 SD and the vegetation index anomaly (VIA) ≤ –1.0 SD. In the copula-transformed space (Fig. 3.3b), this corresponds to the joint probability of $U \leq$ 0.209 and $V \leq$ 0.172, where U and V are the uniform-transformed distributions of SMA and VIA, respectively.

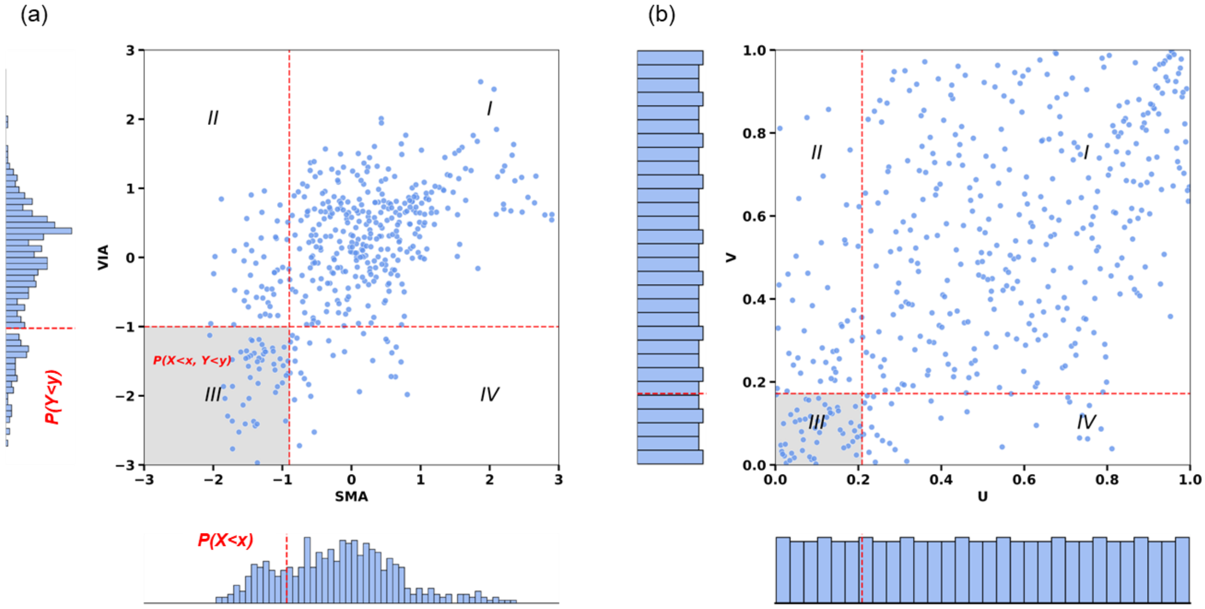


Fig. 3.3. Conceptual illustration of copula transformation from the original bivariate space to the uniform distribution space. The figure demonstrates how joint probabilities between soil moisture anomaly (SMA) and vegetation index anomaly (VIA) are preserved through the copula framework. (a) Bivariate distribution of the original variables: SMA (X) and VIA (Y). (b) Corresponding copula-transformed variables $U = F_X(x)$ and $V = F_Y(y)$, mapped onto the uniform interval [0,1]. Marginal distributions are shown on the axes. The central scatterplots in both panels are divided into four quadrant (I-IV), representing the joint probability regions. The lower-left quadrant (III) corresponds to concurrent low values of both variables (e.g., drought conditions and vegetation deteriorations), illustrating how joint exceedance probabilities are preserved under copula transformation. All quadrant probabilities sum to unity.

3.4.2.3. Understanding vegetation response to soil Moisture deficits: a probabilistic framework

The conditional probability $P(Y \leq y | X \leq x)$, expressed as $P(VIA \leq via |\, SMA \leq sma)$, quantifies the likelihood of vegetation stress occurring under soil moisture deficit conditions. Within the copula framework, this is derived from the joint probability distribution $P(X \leq x, Y \leq y)$, enabling assessment of vegetation deterioration likelihood at various drought severity levels. Formally, the conditional probability is calculated as:

$$P(Y \leq y_j \mid X \leq x_i) = \frac{P(X \leq x_i,\ Y \leq y_j)}{P(X \leq x_i)} = \frac{C[F_X(x_i),\ F_Y(y_j)]}{F_X(x_i)} \quad (3.5a)$$

where X and Y denote the SMA and VIA, and x_i, y_j represent threshold levels in z-score units. For example, the probability of moderate vegetation deterioration (i.e.,$VIA \leq -1.0\ \mathrm{SD}$) under mild soil moisture deficit condition (i.e., $SMA \leq -0.5\ \mathrm{SD}$) is:

$$P(VIA \leq -1.0 \mid SMA \leq -0.5) = \frac{P(SMA \leq -0.5,\ VIA \leq -1.0)}{P(SMA \leq -0.5)} = \frac{F_{SMA,VIA}(-0.5, -1.0)}{F_{SMA}(-0.5)} \quad (3.5b)$$

3.4.2.4. Determination of soil moisture trigger thresholds for vegetation deterioration

To evaluate how vegetation responds to varying levels of soil moisture stress, I derive the conditional cumulative distribution function (CDF) from the fitted copula:

$$F_{Y|X}(y|x) = \frac{\partial C[F_X(x), F_Y(y)]}{\partial F_X(x)} \quad (3.6)$$

where $F_{Y|X}(y|x) = P(Y \leq y | X = x)$, and the right-hand side $\frac{\partial C[F_X(x), F_Y(y)]}{\partial F_X(x)}$ represents the partial derivative of the copula function $C[u, v]$, evaluated at $u = F_X(x)$. This function captures the conditional probability of a VIA (Y) falling below a specified value y, given a known SMA $(X = x)$. The implementation is illustrated in Fig. 3.4. As shown in Fig. 3.4b, the 0.5 and 0.7 probability contours delineate key thresholds where vegetation deterioration becomes increasingly likely under specific SMA conditions. To numerically approximate this function, I adopt a finite-difference method:

$$P(Y \leq y \mid x_2 < X \leq x_1) = \frac{P(x_2 < X \leq x_1, Y \leq y)}{P(x_2 < X \leq x_1)} = \frac{F_{X,Y}(x_1, y) - F_{X,Y}(x_2, y)}{F_X(x_1) - F_X(x_2)} \quad (3.7a)$$

This expression estimates the cumulative conditional probability of vegetation response over a discrete SMA intervals $(x_2,\ x_1]$. Building on this, I develop an iterative threshold search framework to determine the critical SMA level associated with a given vegetation deterioration category. For each interval $(sma - 0.1,\ sma]$, I calculate:

$$P(VIA \leq via_j \mid sma - 0.1 < SMA \leq sma) = \frac{F_{SMA,VIA}(sma, via_j) - F_{SMA,VIA}(sma - 0.1,\ via_j)}{F_{SMA}(sma) - F_{SMA}(sma - 0.1)} \quad (3.7b)$$

Where via_j represents the threshold values of VIA (e.g., $via_j = -1.0\ \mathrm{SD}$ correspond to VD_2 severity level). Starting from sma = 0.0 SD and progressively decreasing in 0.1 SD until the computed cumulative conditional probability exceeds a chosen likelihood decision criterion $P_{decision} \in \{40\%, 50\%, 60\%, 70\%\}$. For instance, if the computed cumulative conditional probability exceeds the chosen likelihood decision criterion $P_{decision}$ (e.g., $P_{decision}$ =60%) at a certain SMA interval, it indicates that there is a 60% chance that this soil moisture deficit will trigger vegetation deterioration at or beyond level via_j. If the probability never exceeds the decision criterion $P_{decision}$ down to sma = -3.0 SD, no effective SMA trigger threshold exists for that vegetation deterioration severity level.

In summary, this methodology enables probabilistic estimation of vegetation deterioration risks across multiple severity levels, based on soil moisture deficit conditions evaluated at cumulative threshold (i.e., $X \leq x$ in Eq. 3.5), exact values (i.e., $X = x$ in Eq. 3.6), and defined intervals (i.e., $x_2 < X \leq x_1$ in Eq. 3.7). By systematically identifying critical SMA triggers, the framework supports improved drought risk assessment and enhances early warning capabilities. The derived SMA thresholds offer actionable insights for proactive drought preparedness across both agricultural and natural ecosystems.

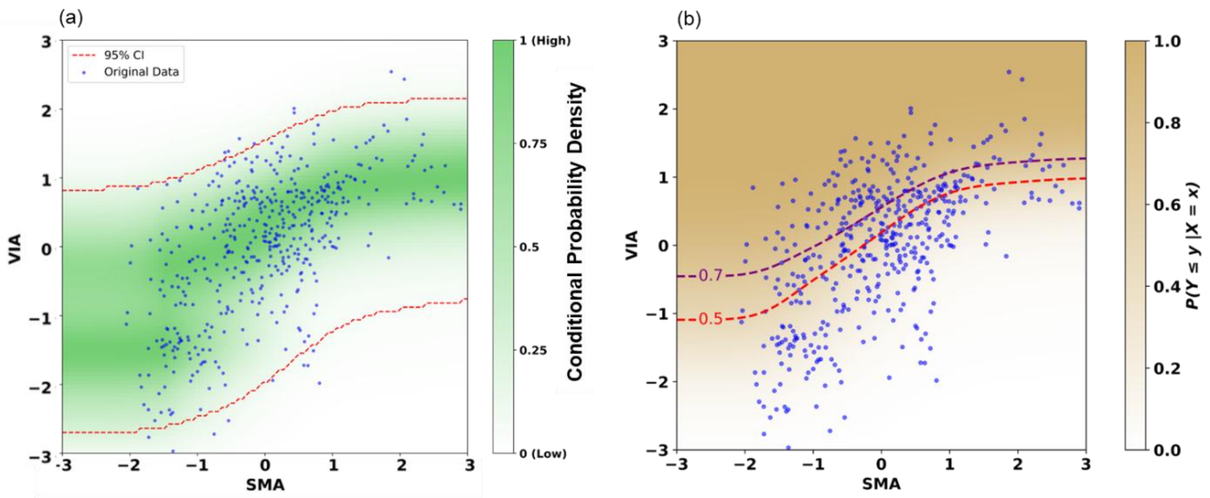


Fig. 3.4. Copula-based conditional probability visualisation at a representative pixel. (a) Conditional probability density function $f(\mathrm{VIA|SMA}) = \frac{\partial}{\partial_y} P(Y \leq y|X = x)$, where the shaded green colour gradient indicates density levels and blue dots represent observed data. The red dashed lines show the 95% confidence interval of the fitted distribution. (b) Cumulative conditional probability $P(Y \leq y|X = x)$. The shaded orange background indicates increasing probability values, with darker tones representing higher cumulative probabilities. The red and orange contours correspond to probability thresholds of 0.5 and 0.7, respectively, highlighting where the likelihood of vegetation deterioration exceeds these levels under given soil moisture deficit conditions.

3.4.3. Predicting vegetation deterioration from soil moisture deficit condition

This section presents a deterministic framework for predicting categorical vegetation deterioration risk (VD_i) using soil moisture deficit conditions. The methodology employs the determined SMA trigger thresholds $(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$ from *Subsection 3.4.2.4* to forecast vegetation deterioration risk across distinct severity levels. For each pixel, I utilised the SMA at time step t to predict the occurrence of vegetation deterioration at specific severity level VD_i (with $i = 1, 2, 3, 4, 5$) at a future time step $t + L$, where L represents pixel-specific temporal lag between VIA and SMA, as quantified in *Subsection 3.4.2.1*.

The framework includes four estimated sets of SMA trigger thresholds, determined through conditional probability decision criteria $P_{decision} \in \{40\%, 50\%, 60\%, 70\%\}$. The prediction mechanism employs an indicator function defined as:

$$\mathrm{VD}_{t+L} = \begin{cases} 0, & \text{if } \mathrm{SMA}_t > \theta_1^{(P_{decision})} \\ i, & \text{if } \mathrm{SMA}_t \leq \theta_i^{(P_{decision})} \end{cases}, (i = 1, 2, 3, 4, 5) \tag{3.8}$$

where $\theta_i^{(P_{decision})}$ represents the trigger threshold value satisfying $P(\mathrm{VD}_i | \mathrm{SMA} \leq \theta_i) \geq P_{decision}$.

For instance, using a decision criterion of $P_{decision} = 60\%$, the most severe deterioration risk (VD_5) is predicted when $\mathrm{SMA}_t \leq \theta_5^{(P_{decision})}$; while moderate risk ($\mathrm{VD}_2$) is indicated when $\theta_3^{(P_{decision})} < \mathrm{SMA}_t \leq \theta_2^{(P_{decision})}$. No vegetation deterioration risk (VD_0) is anticipated when $\mathrm{SMA}_t > \theta_1^{(P_{decision})}$, indicating soil moisture conditions above all critical threshold levels.

To evaluate the predictive capability of this proposed framework, I applied a persistence prediction (PERS) approach as a naïve baseline, following Tian et al. (2019) and Hao et al. (2017), where the *persistence* implies no change in the most recent condition. This baseline assumes that the vegetation deterioration situation at time step $t + L$ remains unchanged from that at time t, i.e., $\mathrm{VD}_{t+L} = \mathrm{VD}_t$.

The prediction sequence is initiated at the beginning of each calendar year, and the performance of VDs prediction is evaluated during the defined peak growing season for each pixel. The modelled VD predictions, generated under different exceedance probability decision criteria $P_{decision} \in \{40\%, 50\%, 60\%, 70\%\}$, are hereafter referred to as Copula-cp40, Copula-cp50, Copula-cp60 and Copula-cp70, respectively. Notably, all these VD prediction sets (Copula-cp40 to -cp70) share the same pixel-wise fitted optimal copula function. The only distinction among them lies in the choice of the exceedance probability decision criterion used to determine the SMA trigger thresholds.

3.4.4 Evaluation metrics for prediction performance

To assess the performance of the proposed framework under different configurations, I conducted a comparative evaluation against persistence prediction, which serves as a benchmark. Given the dual challenges of ordinal multi-categorical classification and imbalanced sample sizes in the VD risk classification system, I employed a suite of complementary evaluation metrics to comprehensively quantify predictive skill. These include the multi-category (6 × 6) contingency table with hit-miss statistics (Table 3.1), the Root Mean Square Error (RMSE), and the frequency distribution analysis approach. These metrics quantify the agreement between predicted VD categories (y_i) and observed categories (o_i). The metrics can be calculated overall and for each category.

3.4.4.1. Hit-miss statistic metrics

The contingency-based metrics, including Hit Rate (Probability of Detection), False Alarm Ratio (FAR), and Overestimation and Underestimation Ratios, capture different aspects of model performance. They are particularly useful in diagnosing errors across rare and extreme events, where conventional accuracy measures may fall short. These definitions and their applications are adapted from widely used frameworks in environmental prediction and risk assessment literature (e.g., Bennett *et al.,* 2013; Papagiannaki *et al.,* 2020; Mardian *et al.,* 2023). These studies offer precedent for using hit-miss statistics to evaluate probabilistic or multi-category forecasts, especially under skewed distributions. The inclusion of directional error metrics (over-/underestimation) is especially relevant for early warning systems, where missed warnings (underestimation) or false alarms (overestimation) have markedly different implications for decision-making.

Table 3.1. Hit-miss metrics used to assess prediction performance of vegetation deterioration categories.

Metric name	Formula	Range	Ideal value	Notes
Hit Rate (Probability of Detection, PoD)	$\frac{\text{hits}}{\text{hits} + \text{misses}}$	(0,1)	1	Sensitive to hits, but ignores false alarms. Good for rare events
False Alarm Ratio (FAR)	$\frac{\text{false alarms}}{\text{hits} + \text{false alarms}}$	(0,1)	0	Sensitive to false alarms, but ignores misses
Overestimation Ratio	$\frac{\text{misses}_i \; where \; (\text{y}_i > o_i)}{\text{hits}_i + \text{misses}_i}$	(0, 100%)	0	Useful in quantifying conservative prediction bias.
Underestimation Ratio	$\frac{\text{misses}_i \; where \; (\text{y}_i < o_i)}{\text{hits}_i + \text{misses}_i}$	(0, 100%)	0	Crucial for evaluating the missed warning signals.

3.4.4.2. RMSE

To assess the overall prediction performance of my model, I employed the categorical averaged Root Mean Square Error (RMSE), defined as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_i - O_i)^2} \tag{3.9}$$

where Y_i and O_i represent the predicted and observed VD categories at timestamp i, respectively, and *N* denotes the total number of prediction repeats. The RMSE quantifies the average magnitude of prediction errors, with lower values indicating greater accuracy. This metric evaluates the level of agreement between predictions and observations, though it does not indicate directional bias in the predictions.

3.4.4.3 Frequency distribution

I utilised 6 × 6 contingency tables to analyse the frequency distribution of predicted versus observed frequencies per vegetation deterioration (VD) category. The contingency tables facilitate analysis of estimation errors through examination of off-diagonal frequencies, distinguishing between overestimation and underestimation patterns. This approach enables detailed ability examination of each estimated set of SMA trigger thresholds to predict specific VD levels. The performance evaluation incorporates multiple verification metrics, namely the Hit Rate (or POD), False Alarm Ratio (FAR), and the Over- or Underestimation Ratio.

3.5. Results

3.5.1. Dependencies between root-zone soil moisture and vegetation dynamics

The analysis of regionally averaged time series of VIA and SMA revealed strong temporal coupling between vegetation growth dynamics and root-zone soil moisture conditions. As shown in Fig. 3.5, vegetation growth dynamics were closely related to root-zone soil moisture conditions, with general good agreement observed throughout the study area. Intense soil moisture droughts, particularly the 2006 drought event (dark brown shaded band in Fig. 3.5), corresponded with notable negative durations of VIA. All vegetation deterioration episodes, characterised by negative VIA values during the study period, coincided with sustained root-zone soil moisture deficits. The region experienced significant drought events in 2002, 2005, 2006, and 2018.

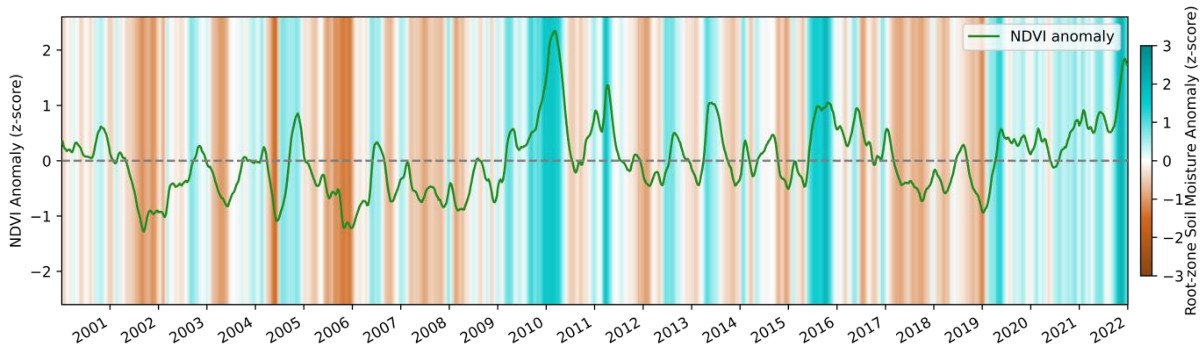


Fig. 3.5. The one-month rolling mean standardised anomalies of the NDVI (2001-2022 climatology). Background shading represents one-month rolling mean standardised root-zone soil moisture anomalies.

The Fig. 3.6a illustrates the maximum Spearman's rank correlation coefficients (ρ_{max_lag}) between the SMA and VIA in each pixel that passed the significance test (p<0.01). Most of the study area exhibited positive correlations, with shallow-rooted crop and pasture lands in central and northeastern regions showing strong lagged effects (ρ_{max_lag} > 0.6). The ρ_{max_lag} values were concentrated between 0.4 and 0.8, while negative SMA-VIA correlations were mainly observed in western semi-arid shrublands and eastern alpine forest areas (Fig. 3.6b). The temporal response patterns, depicted in Fig.3.6c, show vegetation response to root-zone soil moisture is primarily concentrated within 8 to 40 days across the study area, where a positive lag indicates that SMA leads VIA. Regions with response times less than one month comprise approximately 65% of the study region, typically coinciding with relatively high ρ_{max_lag} values. Notably, dryland cropping areas exhibit longer response lag times exceeding one month (Fig. 3.6d). Similar extended response lags were observed in southeastern alpine forests and northwestern shrublands, although the SMA-VIA relationship in these regions generally persisted only over short temporal lags (around 2-4 of the 8-day intervals). A distinct pattern emerged in western natural environments, characterised by sandplain acacia shrublands, where negative correlations indicated vegetation changes preceded soil moisture variations in the root-zone layer (0-1.0m), contrasting with the predominant lagged response pattern observed across other regions.

The spatial distribution of optimal copula functions, selected based on the K-S test, CvM test and AIC criteria, revealed distinct dependency patterns across different land use types (Fig. 3.7a). In modified pasture regions, the Clayton copula predominated (Fig. 3.7b), a family known for capturing stronger lower tail dependence (i.e., joint extreme low values) (Nelsen, 2006). This pattern implies heightened vulnerability of modified pastures to concurrent soil moisture deficits and vegetation deterioration. Dry cropping areas, in contrast, exhibited better fit with the Gumbel copula, which is characterized by stronger upper tail dependence (Nelsen, 2006). This characteristic reflects a strong positive association between soil moisture and vegetation health under favourable conditions in dry cropping lands. Native pasture areas displayed a distinct dependence structure, predominantly characterised by Frank copulas. This pattern indicates a relationship between SMA and VIA that is symmetric with respect to the tails, lacking the strong lower or upper tail dependence observed in modified pasture and dry cropping lands.

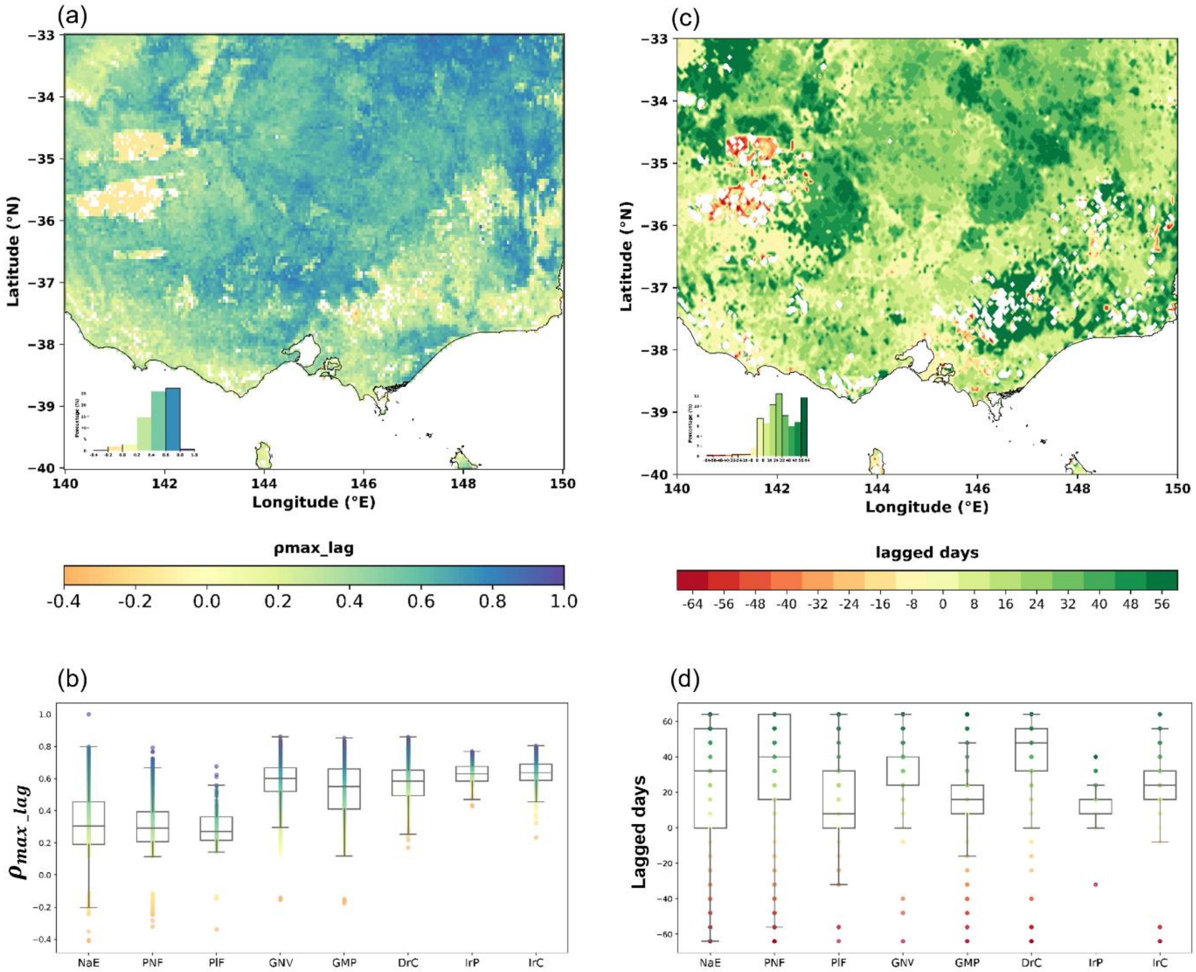


Fig. 3.6. Spatial patterns and land-use differences in the correlation and response time between soil moisture anomaly (SMA) and vegetation index anomaly (VIA) during the growing seasons (2001–2022). (a) Maximum Spearman's rank correlation coefficients ($\rho_{max\ \_lag}$) at each pixel, showing the strength of SMA-VIA association. (b) Boxplot of $\rho_{max\ \_lag}$ grouped by land-use types. (c) Lagged response time (in 8-day interval)) at which maximum correlation occurs, mapped only for statistically significant pixels (p < 0.01). (d) Boxplot of lagged response days by land-use types. Land-use types: NaE – Nature conservation; PNF – Production native forest; PlF – Plantation forestry; GNV – Grazing native vegetation; GMP – Grazing modified pasture; DrC – Dryland cropping; IrP – Irrigated pasture; IrC – Irrigated cropping.

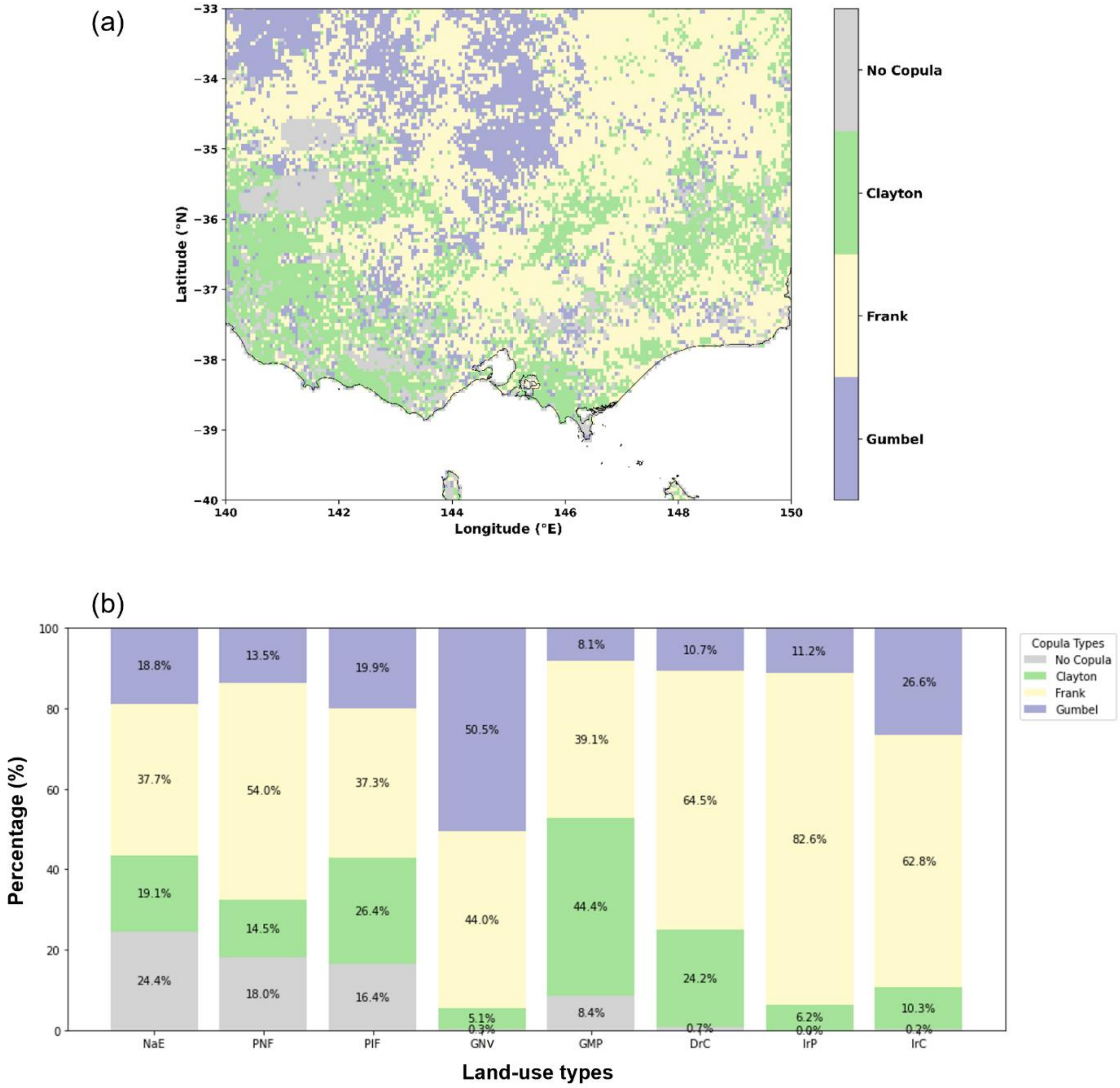


Fig. 3.7. Spatial and categorical distribution of best-fitted copula types for modelling the joint relationship between soil moisture anomaly (SMA) and vegetation index anomaly (VIA) across southeastern Australia. (a) Spatial distribution of selected copula families at each pixel. Clayton, Frank, and Gumbel copulas are colour-coded as shown. Grey indicates pixels where no suitable copula was fitted (non-significant dependence), and white areas denote water bodies or regions without vegetation cover. (b) Proportional distribution of copula types across land cover classes. The bars represent the relative frequency (%) of each copula family fitted within different vegetation and land-use types. Land-use types: NaE – Nature conservation; PNF – Production native forest; PIF – Plantation forestry; GNV – Grazing native vegetation; GMP – Grazing modified pasture; DrC – Dryland cropping; IrP – Irrigated pasture; IrC – Irrigated cropping.

3.5.2. Conditional probabilities of vegetation deteriorating under soil moisture drought scenarios

I employed copula theory to quantify the probability of vegetation deterioration under varying soil moisture drought scenarios. Fig. 3.8 illustrates the conditional probabilities of different vegetation deterioration severity levels under five soil moisture drought scenarios (mild, moderate, severe, extreme, and exceptional). The colour intensity within each pixel represents the probability of vegetation deterioration under corresponding drought conditions. For any given vegetation deterioration threshold (Fig. 3.8, row panels), both conditional probability and affected area increased with intensifying soil moisture stress (SMA values from -0.5 to -2.0 SD). The likelihood of VIA falling below -0.5 SD increased markedly as moderate soil moisture drought progressed to severe or extreme conditions. This pattern persisted for more severe vegetation deterioration thresholds (VIA ≤ -1.0 and -1.5 SD).

Spatial analysis reveals that central and northeastern regions exhibited higher probabilities of vegetation deterioration during the growing season, consistent with the spatial pattern of SMA-VIA correlation strength (Fig. 3.6a). Semi-arid areas with high interannual hydrological variation show greater risk of vegetation deterioration compared to arid and temperate climate zones. This enhanced vegetation-soil moisture dependency in semi-arid areas, particularly in managed land use types, indicates that soil moisture drought was the primary driver of vegetation deterioration. At the regional scale, the average conditional probability of severe vegetation deterioration (VIA ≤ -1.5 SD) was 14.5% under mild soil moisture drought (SMA ≤ -0.5 SD), increasing significantly to 19.1%, 25.4%, 29.8%, and 31.8% under moderate, severe, extreme, and exceptional soil moisture drought scenarios, respectively (see Supplementary Fig. S3.1). For more severe vegetation deterioration (VIA ≤ -2.5 SD), the average conditional probabilities were 2.0%, 2.8%, 4.4%, 7.7% and 12.6% under progressively worsening soil moisture drought scenarios.

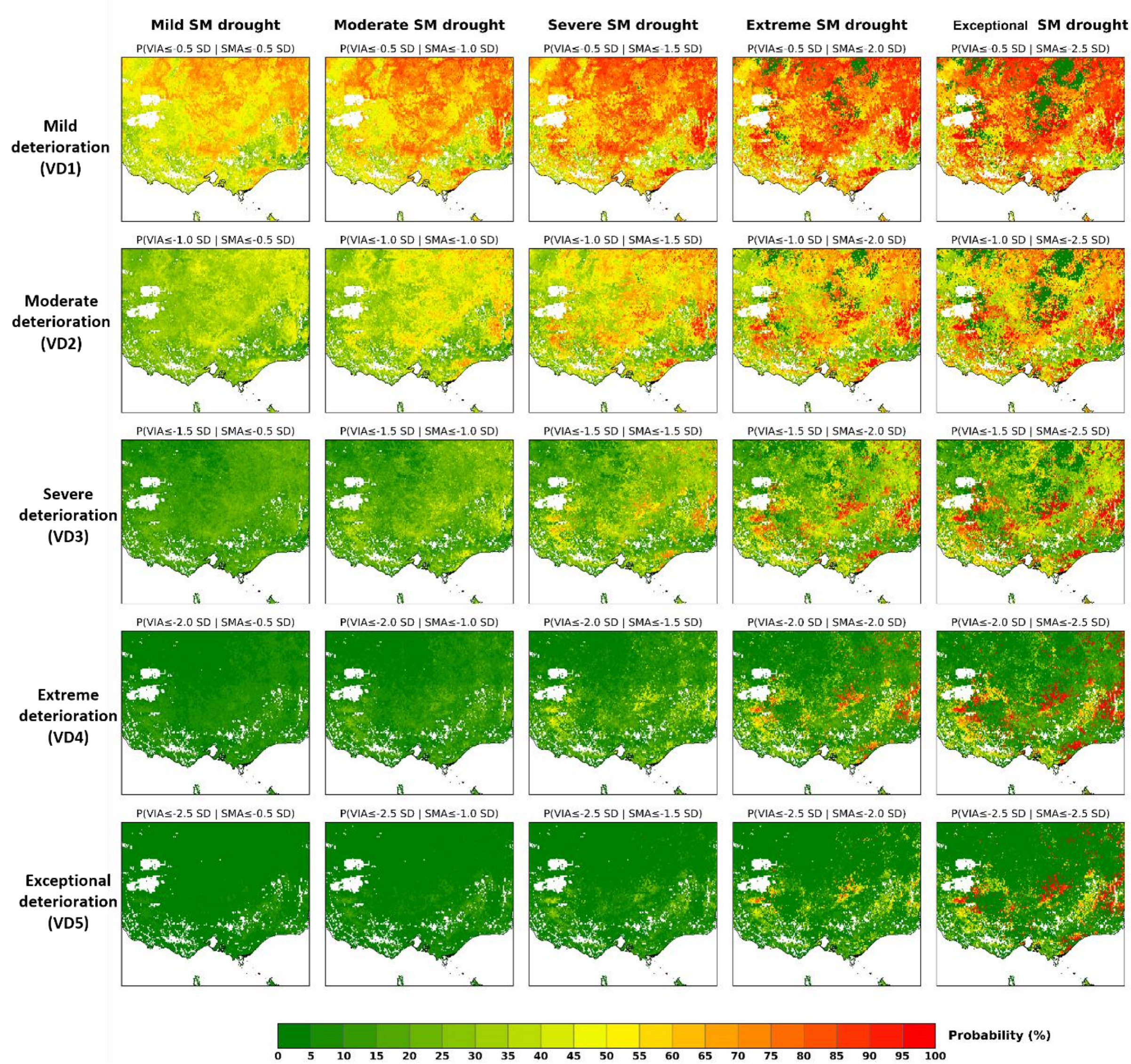


Fig. 3.8. The conditional probabilities of vegetation deterioration under soil moisture drought scenarios (mild, moderate, severe, extreme, and exceptional) during the growing season. Columns show soil moisture drought scenarios and rows show vegetation deterioration severity levels (mild to exceptional). Probability values are represented by the colour intensity within each pixel, with redder colours indicating higher probabilities of vegetation deterioration under the given soil moisture conditions. White areas indicate regions where no corresponding SMA-VIA simulation pairs exist within the specified thresholds or non-vegetated areas.

3.5.3. Critical soil moisture trigger thresholds for vegetation deterioration

To further verify the trigger threshold identifying framework, I examined conditional probabilities of VD occurrence under varying soil moisture stress conditions at four randomly selected pixels (Fig. 3.9). The results confirm that, for any certain VD level (e.g., VIA ≤ -1.5 SD), the probability of occurrence increases monotonically as SMA declines from 0.0 to -3.0 SD. Conversely, for a given SMA level (e.g., SMA at -0.9 SD), the probability of occurrence systematically decreases as the VD threshold becomes more severe (e.g., from VIA ≤ -1.0 SD to ≤ -2.0 SD). The black dashed line in each panel marks the decision criterion $P_{decision}$ =50% for demonstration. These findings substantiate the robustness of the trigger threshold framework established in this study for developing VD early warning systems responsive to varying levels of soil moisture drought.

To characterise regional patterns in the identified least severe soil moisture deficits sufficient to trigger vegetation deterioration events, Fig. 3.10 presents the spatial distribution of critical SMA thresholds for five VD severity levels using a 40% decision criterion ($P_{decision}$=40%) as an illustrative example. The four decision criteria (40-70%) examined in this study to evaluate how trigger performance varies across the selection of probability levels. Each panel displays the minimum SMA value at which a specific VD level is predicted to occur with at least 40% probability. Each panel displays the SMA value required to trigger a given VD category with at least 40% occurrence probability. Darker red shades indicate that more intense soil moisture deficits are required to trigger deterioration, while lighter yellow tones represent areas where relatively mild deficits are sufficient. The proportion of pixels with identifiable SMA thresholds declines with increasing VD severity, because for many locations the conditional probability of severe deterioration never reaches the selected decision criterion (e.g., $P_{decision}$=40%) even at the lowest SMA values (i.e., SMA at -3.0 SD) represented in the fitted copula models. Under this probability criterion, the SMA thresholds were identifiable for 85%, 65%, 29%, 18%, and 14% of the copula-fitted pixels for VD_1 through VD_5, respectively. Central and northern regions exhibit relatively mild trigger thresholds, suggesting heightened sensitivity to soil moisture stress. In contrast, coastal forests and grazing-modified pastures require more severe deficits to trigger the same deterioration levels, reflecting greater resilience.

The violin plot (Fig 3.10f) summarises the distribution of identified SMA thresholds across VD severity levels. Each VD class demonstrates considerable variability in SMA threshold values around the central tendency. For mild (VD_1) and moderate (VD_2) deterioration categories, approximately 75% of the SMA trigger thresholds are concentrated in the lower quartiles. In contrast, VD_3 to VD_5 display more uniform distributions of SMA trigger thresholds across their respective ranges, suggesting a more complicated SMA-VD relationship response across different land use types at higher severity levels. As presented in Supplementary Table S3.3, the median of SMA trigger thresholds become progressively more negative with increasing VD severity: -0.77 SD (VD1, mild), -1.46 SD (VD2, moderate), -2.08 SD (VD3, severe), -2.30 SD (VD4, extreme), and below -2.48 SD (VD5, exceptional).

While the primary analysis in this section focuses on the 40% exceedance probability, additional estimated sets of SMA trigger thresholds with $P_{decision} \in \{50\%, 60\%, 70\%\}$ are presented in the Supplementary Fig. S3.2. These maps reveal a systematic contraction in spatial coverage as the required exceedance probability increases, indicating that fewer pixels meet the higher certain certainty requirement, especially in inland regions. Concurrently, the trigger thresholds shift towards more negative SMA values. For instance, thresholds associated with $P_{decision}$= 70% are typically 0.5 to 1.0 SD lower than those at $P_{decision}$ =40%, reflecting the greater soil moisture deficit needed to trigger vegetation deterioration with higher confidence.

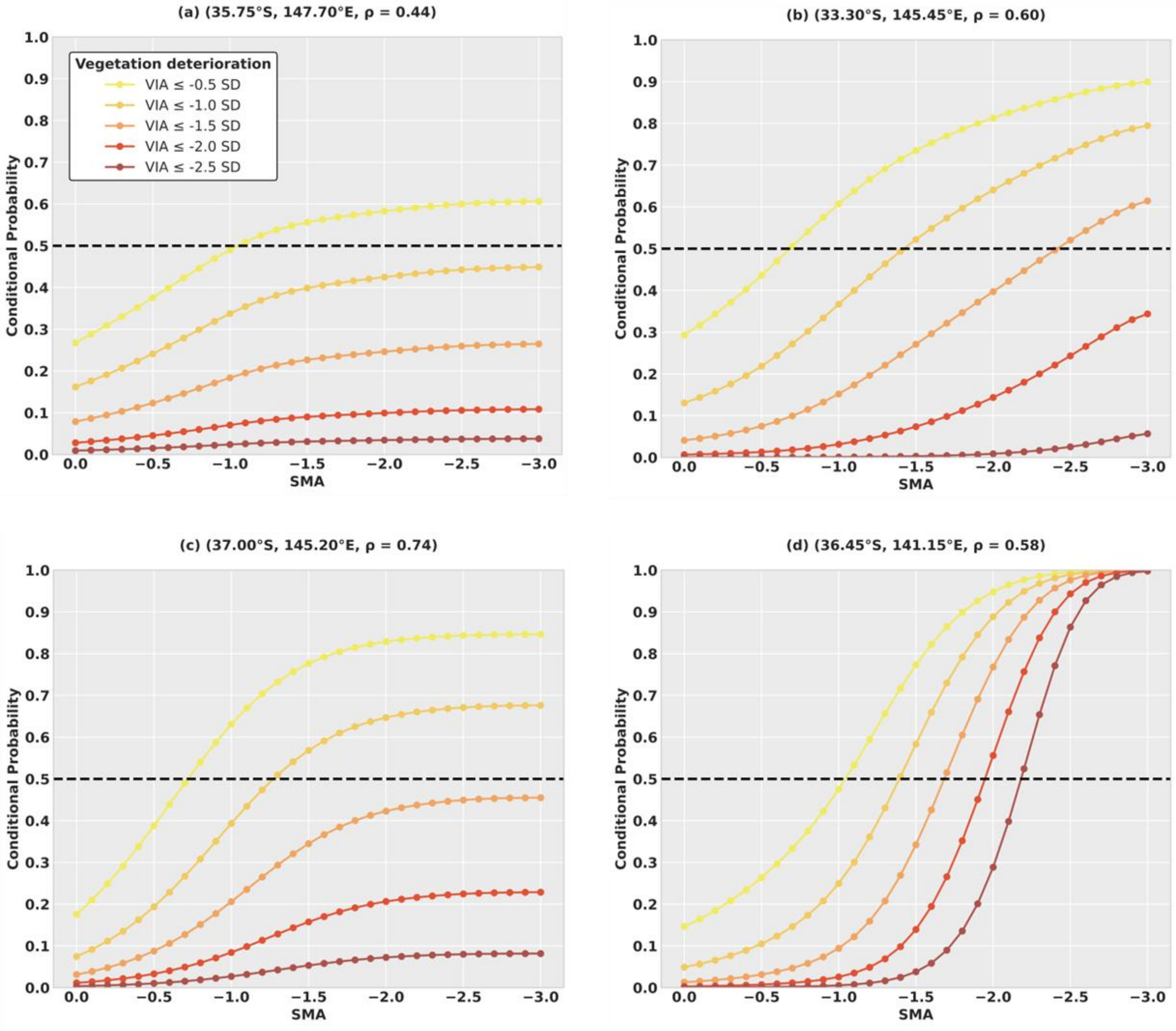


Fig. 3.9. Conditional probability curves at four randomly selected pixels across major land cover types, illustrating the variation in vegetation deterioration (VD) probabilities under varying soil moisture anomalies (SMA). Each panel represents a distinct land cover: (a) plantation forest, (b) grazing native vegetation, (c) grazing modified pasture, and (d) dryland cropping. Coloured lines correspond to five severity levels of VIA. Black dashed lines indicate the chosen exceedance probability criterion ($P_{decision}$=50%). Observed maximum correlation coefficients (ρ) of SMA-VIA pairs in the pixels are also reported.

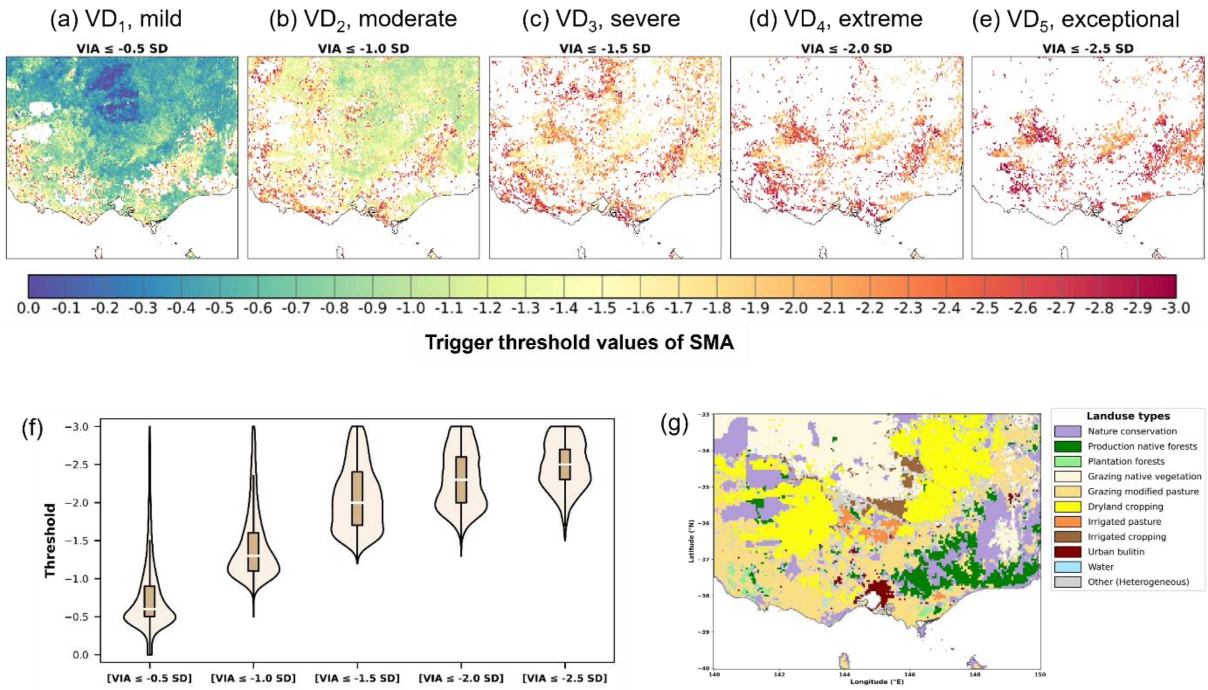


Fig. 3.10. Spatial distribution and statistical summary of SMA trigger thresholds for VD categories based on a 40% exceedance probability criterion ($P_{decision}$=40%). Panels (a-e) illustrate the spatial patterns of SMA trigger thresholds corresponding to increasing vegetation deterioration severity: (a) mild (VIA ≤ -0.5 SD), (b) moderate (VIA ≤ -1.0 SD), (c) severe (VIA ≤ -1.5 SD), (d) extreme (VIA ≤ -2.0 SD), and (e) exceptional (VIA ≤ -2.5 SD) conditions. Colour gradients indicate the severity of required soil moisture deficit conditions, with dark red shades representing more negative SMA values necessary to trigger corresponding VD events. White areas indicate pixels without determined SMA thresholds due to model fitting constraints. Panel (f) presents violine plot showing the distribution of SMA thresholds across VD categories. The upper and lower limits of the SMA trigger threshold are 0.0 and -3.0 SD, respectively. Subplot (g) presents the land use types of the study region.

3.5.4. Evaluation of predictive performance and model skill

3.5.4.1. Overall accuracy

Table 3.2 presents the average verification metrics derived from pixel-level 6 × 6 contingency tables, quantifying the agreement between observed and predicted vegetation deterioration events for different categories (no-risk to exceptional, VD_0 to VD_5) across all vegetated land-use types. Results are reported for the four estimated sets of SMA thresholds across southeastern Australia (Copula-cp40, Copula-cp50, Copula-cp60, Copula-cp70), alongside the persistence prediction (PERS) serving as the benchmark.

The verification metrics indicate that copula-based models tend to underestimate VD risks more frequently than the persistence prediction benchmark, with underestimation rates ranging from 73% to 82% at the 8-day temporal resolution (versus 42% for persistence), and 68% to 81% at the monthly aggregation period (versus 36% for persistence). The monthly aggregation was achieved by comparing the maximum observed VD event in each month with the maximum predicted VD event for that same period. These results highlight the inherent difficulty of probabilistic multi-category forecasting, especially for infrequent high-severity events. Nonetheless, among the copula-based configurations, Copula-cp40 demonstrates relatively stable skill across categories and timescales, achieving the highest overall Hit Rates (0.23 and 0.25) and the lowest FAR (0.67 and 0.64), suggesting its potential value in risk-aware forecasting frameworks that offer category-specific probabilistic forecasts rather than relying solely on temporal continuity. The full contingency tables of each VD risk category are presented in the following section (Section 3.5.4.2).

Although the copula-based configurations yield higher overall RMSEs compared to the persistence prediction benchmark (0.98–1.09 vs. 0.71 at 8-day scale; 1.07–1.30 vs. 0.73 at monthly scale), spatial disaggregation of RMSE differences (Fig. 3.11) provides further insights into where copula-based models may offer added value. In particular, Copula-cp40 demonstrates improved performance in several regions dominated by forests and grazing-modified pastures, as indicated by the blue shading in Fig 3.11, where it achieves lover RSM than the PERS benchmark. These regional improvements should not be overinterpreted as universal gains, but rather as indicative of landscape-specific advantages, especially in heterogeneous environments where persistence-based methods may struggle to adapt. Conversely, in more homogeneous cropping areas, the models perform comparably, reflecting the limited incremental benefit under stable dynamics.

Table 3.2. Evaluation and verification of estimated sets of SMA thresholds (Copula-cp40 to Copula-cp70) and persistence prediction (PERS) at 8-day and monthly timescales. The values represent average performance across all pixels with vegetated land-use types, based on statistics from pixel-level 6 × 6 contingency tables covering the full spectrum of vegetation deterioration categories (no-risk to exceptional, VD_0 to VD_5).

	Verification measures (at 8-day timescale)				
Sets of predictions	RMSE	Hit Rate	FAR	Under-estimation (%)	Over-estimation (%)
Copula-cp40	0.98	0.23	0.67	73	4
Copula-cp50	1.02	0.21	0.67	78	2
Copula-cp60	1.06	0.19	0.68	81	1
Copula-cp70	1.09	0.18	0.69	82	0
Persistence prediction (PERS)	0.71	0.48	0.49	42	10

Note: RMSE (perfect=0), Hit Rate: (perfect = 1), FAR: False Alarm Ratio (perfect = 0)

	Verification measures (at monthly maximum aggregation)				
Sets of predictions	RMSE	Hit Rate	FAR	Under-estimation (%)	Over-estimation (%)
Copula-cp40	1.07	0.25	0.64	68	7
Copula-cp50	1.14	0.23	0.64	74	3
Copula-cp60	1.22	0.20	0.66	79	1
Copula-cp70	1.30	0.18	0.67	81	1
Persistence prediction (PERS)	0.73	0.54	0.43	36	10

Note: RMSE (perfect=0), Hit Rate: (perfect = 1), FAR: False Alarm Ratio (perfect = 0)

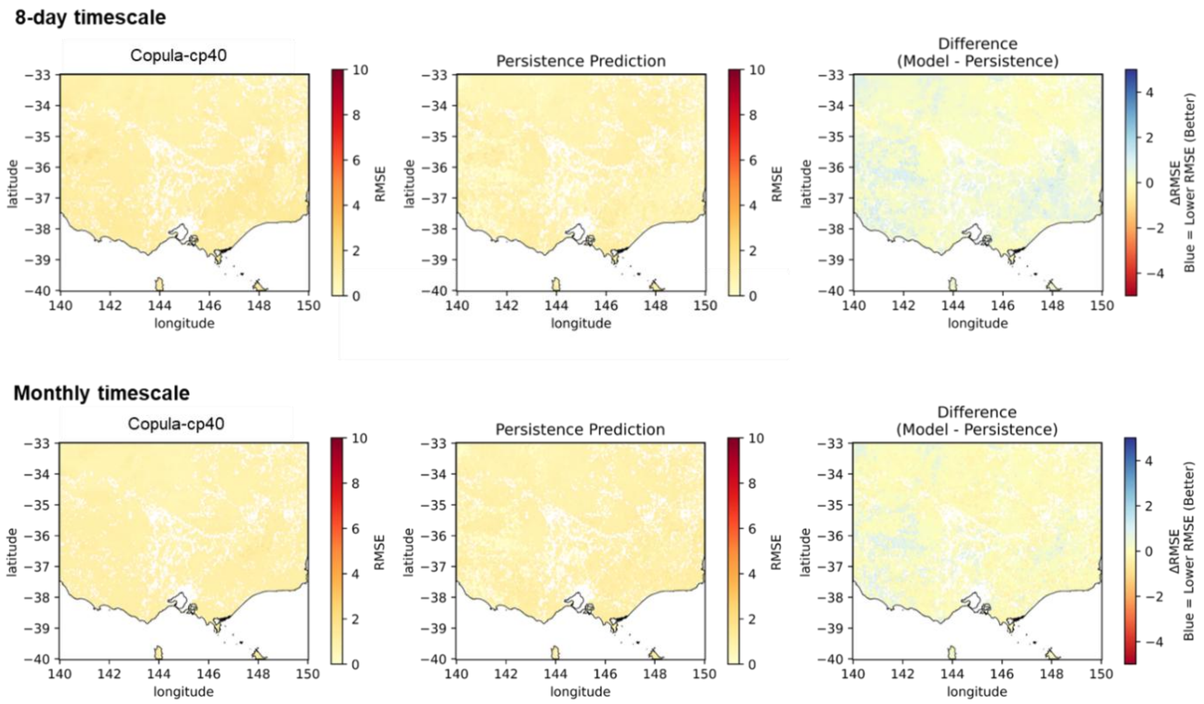


Fig. 3.11. Spatial comparison of RMSE between the Copula-cp40 VDs predictions and the persistence prediction (PERS) benchmark at 8-day (top row) and monthly maximum aggregation (bottom row) timescales. The difference panels (right column) illustrate RMSE improvements, where blue shading indicates areas where Copula-cp40 outperforms the PERS benchmark (i.e., lower RMSE), particularly across heterogeneous landscapes such as forested and pasture regions.

3.5.4.2 Performance of estimated SMA thresholds per VD risk category

Fig. 3.12 illustrates the relative frequency distributions (%) of vegetation deterioration events across VD categories for both copula-based predictions and persistence forecasting (PERS). The black-outlined bars represent correct classifications (hits), while the adjacent bars indicate misclassifications into other VD categories (misses), demonstrating the model predictive capabilities across the severity grades. Analysis of the frequency distributions reveals several key patterns in predictive performance:

- Model-specific analysis reveals that Copula-cp40 demonstrates the most balanced performance across categories. Increasing conditional probability thresholds (cp50-cp70) shows progressive enhancement in no-risk detection but at the cost of increased underestimation in higher risk categories. PERS maintains superior performance in extreme categories but shows less consistency in lower risk classifications.

- For the performance in lower risks, the copula-based models demonstrate comparable or marginally improved hit rates relative to PERS for the no-risk (VD_0) and mild (VD_1) categories. Specifically, the hit rate for no-risk conditions reaches 0.99 with Copula-cp70, compared to 0.88 for PERS. The detection of mild deterioration events shows notable variation, with Copula-cp40 achieving a hit rate of 0.33, while PERS maintains a rate of 0.45.
- For moderate risk events (VD_2), detection capability declines markedly across all copula-based models, with hit rates diminishing from 0.14 (Copula-cp40) to 0.01 (Copula-cp70), compared to PERS (0.41).
- In contrast, higher severity categories (VD_3-VD_5) show pronounced underestimation, particularly in copula variants with setting of the higher conditional probability criterion.

These findings indicate that whilst copula-based models offer competitive performance for lower risk categories, they display systematic limitations in capturing severe vegetation deterioration events. This behavioural pattern reveals a trade-off between reliable detection of normal conditions and accurate identification of extreme events. The consistent underestimation in higher risk categories impairs the overall capability. Nevertheless, the low overestimation rates indicate minimal false alarm potential, suggesting operational utility in risk assessment frameworks despite their reduced sensitivity to extreme vegetation deterioration events.

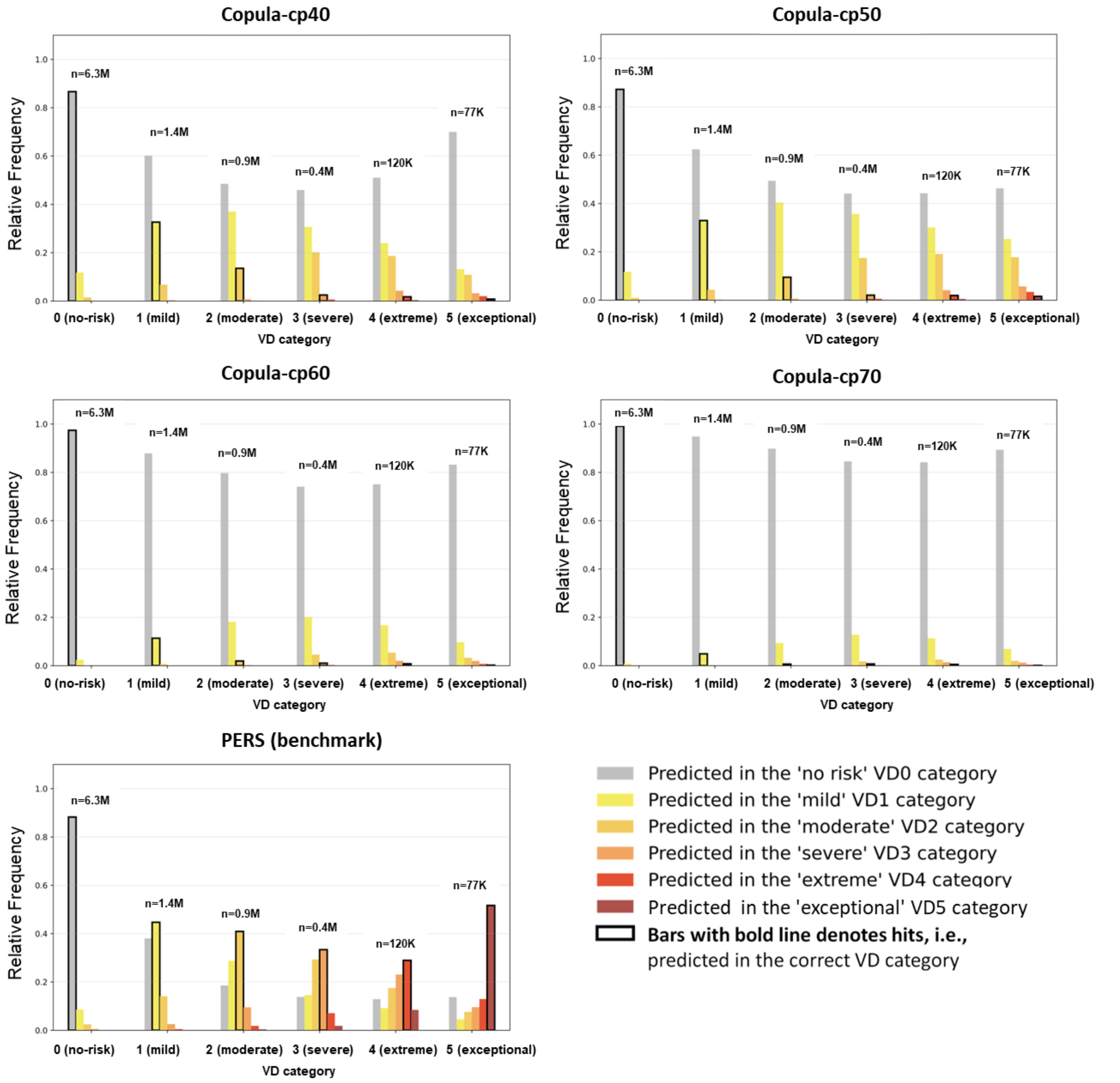


Fig. 3.12. Distribution of predicted vegetation deterioration (VD) risk categories relative to observed categories on an 8-day temporal resolution. Each panel shows relative Hit/Miss frequency for different models: four copula-based models with varying conditional probability thresholds (cp40, cp50, cp60, and cp70) and the persistence prediction (PERS) as reference. Bars represent the predicted frequency in each VD category. The black-outlined bars indicate correct predictions (hits) for each category. The sample sizes (n) are provided for each VD category. Due to large values, they are expressed in thousands (K) or millions (M) for clarity. Monthly scale distributions are presented in Supplementary Material Fig. S3.3.

3.6. Discussion

3.6.1. Methodological limitations and uncertainties

While the probabilistic trigger threshold framework demonstrates utility for predicting vegetation deterioration based on root-zone soil moisture deficits, several limitations should be considered. First, the framework focuses solely on soil moisture deficit as the driver of vegetation deterioration, whereas other factors such as insect infestations and fire disturbances are not accounted for in the threshold identification. Although justified by the strong SMA-VIA correlations across the study region (Fig. 3.2a), this focus may oversimplify the broader ecological processes contributing to vegetation decline. Consequently, the identified thresholds should be interpreted as capturing a primary hydrological signal rather than a comprehensive assessment of all ecological stressors.

The physical credibility of soil moisture estimates is supported by extensive validation of the AWRA-L product against in-situ observations (Frost *et al.,* 2021; Holgate *et al.,* 2016; Vogel *et al.,* 2021). However, the standardisation of these anomalies relies on a 22-year baseline (2001–2022). While reflecting contemporary conditions, this period may be insufficient to capture full decadal variability or to derive stable anomaly distributions in transitional climates. Since shifting the baseline period (i.e., the temporal window used to define normal conditions) can markedly alter the frequency and severity of detected anomalies, the trigger thresholds should be interpreted with caution in regions where longer-term climatological records might yield different results.

A further limitation is the use of a static snapshot for 2016 land-use map (ACLUMP v8) to represent the entire 2001-2022 period introduces potential uncertainty. While broad land-use transitions across the study region are relatively limited compared with the magnitude of climatic variability, this simplification may still affect the characterisation of vegetation responses in areas experiencing changes in land utilisation.

3.6.2. Challenges in predicting extreme events

My analysis highlights the inherent difficulty in predicting extreme VD events, as shown by the marked decline in predictive skill as severity increases (Table 3.2). Hit rates for the copula-based models diminished from 0.33 for mild categories to below 0.02 for extreme categories, a degradation primarily driven by two factors.

Firstly, the relatively short observational period (2001-2022) provides limited sampling of extreme events, particularly for severe categories (VD_3~VD_5). The scarcity of extreme events in observation-based data has been noted in previous studies (Sun *et al.,* 2023a; Yang *et al.,* 2023a; Zhang *et al.,* 2023). Whilst satellite remote-sensing products offer extensive spatiotemporal coverage of vegetation conditions, they nevertheless capture few drought-induced deterioration events at monthly or finer temporal scales. This limited sampling reflects a fundamental constraint of empirical approaches: the rarity of extreme events restricts the robustness of coincidence analysis with the practical inability to observe sufficient extreme events for another decades.

Secondly, even though copula-based approaches excel theoretically at modelling tail dependencies, which is particularly relevant for analysing both linear and non-linear soil moisture-vegetation relationships across different land use types (Sun *et al.,* 2023), their practical implementation presents several challenges. The selection of marginal distributions and copula functions, though offering flexibility, introduces uncertainties that propagate through the modelling chain, affecting conditional probability estimates and subsequent trigger threshold determinations (Achite *et al.,* 2022). As evidenced in Fig. 3.10, extreme events (VD_4) display greater spatial heterogeneity in their conditional probabilities compared to moderate conditions (VD_2). This suggests more complex underlying relationships that my bivariate copula framework may not fully capture, particularly across different land use types. Therefore, the interaction between land use characteristics, copula function selection, and prediction skill reveals systematic patterns that warrant further investigation for improving extreme event prediction capabilities. This complexity in extreme event prediction highlights the need for more sophisticated approaches that can better account for the non-linear relationships and spatial heterogeneity inherent in soil moisture-vegetation dynamics during extreme drought conditions.

3.6.3. Threshold selection and copula-based framework merits

My analysis confirms that drought occurrence does not necessarily lead to proportional vegetation deterioration. For southeastern Australia's growing season, even under severe soil moisture deficit conditions (SMA ≤ -2.5 SD), the region-wide probability of experiencing mild vegetation deterioration events averages only 61.2% (Fig. S3.1). This finding aligns with previous research on dryland vegetation responses, which has highlighted the complex and nonlinear relationship between hydrological stress and ecosystem impacts (X. Li *et al.,* 2023; Van Loon & Laaha, 2015; Zhang *et al.,* 2022b).

While a 50% conditional probability criterion is commonly adopted to indicate an even chance of occurrence, my evaluation identified superior performance using a 40% criterion (i.e., Copula-cp40). In operational terms, lower decision thresholds such as 40% favour a more conservative early-warning posture, prioritising the reduction of missed alarms at the expense of higher false-alarm rates. This finding reflects the relatively low coincidence probabilities observed across my study region. As shown in Fig. 3.10a–e, increasing the conditional probability threshold results in more conservative trigger identification but reduces spatial coverage of valid predictions, illustrating the trade-off between false alarm reduction and early warning inclusivity. Temporal resolution also influenced model performance, with monthly-scale predictions (aggregated using monthly maxima) slightly outperforming 8-day predictions (Fig 3.11; table 3.2).

The temporal relationship patterns (Fig 3.6c) underscore that the effective temporal window for forecasting vegetation deterioration from 'near-real-time' soil moisture signals is relatively short, typically within 0-1 month, particularly in short-rooted systems such as dryland cropping and grazing-modified pastures. All copula-based models systematically underestimated deterioration risks compared to the persistence prediction (PERS) benchmark, an outcome expected given that the benchmark directly extrapolates prior vegetation states. Notably, the copula-based approach performed more reliably in predicting no-risk to moderate-risk conditions, while skill declined for higher-risk categories. The superior performance of the persistence prediction benchmark for short-term prediction aligns with previous studies using autoregressive models and Gaussian processes (Barrett et al., 2020), linear regression (Zambrano et al., 2018) and Bayesian autoregression (Salakpi et al., 2022), as the persistence prediction relying solely on historical vegetation index values to predict future states. While effective at exploiting temporal continuity, these approaches are inherently limited in their ability to incorporate hydro-meteorological forecasts or to capture multivariate drought drivers. In contrast, the copula-based framework provides a probabilistically explicit linkage between root-zone soil moisture and vegetation deterioration risk categories. This structure facilitates integration with forecast soil moisture data, enabling forward-looking and risk-based early warning applications. Although current model performance is constrained by category imbalance and limited sample sizes for extreme events, the framework offers a scalable and interpretable foundation for advancing drought impact forecasting systems beyond single-variable persistence logic.

3.6.4. Future research directions

My bivariate copula approach has demonstrated utility for modelling moderate vegetation stress conditions. Future developments could explore the potential of incorporating additional climate variables (such as temperature and evaporation) into a multivariate copula framework to improve sensitivity to compound drought stressors. However, such extensions would need to carefully address the small sample limitations inherent in higher-dimensional modelling, particularly for severe vegetation deterioration categories where observational records are sparse. In parallel, using alternative vegetation indices beyond NDVI, such as the Enhanced Vegetation Index (EVI) or Solar-induced Chlorophyll Fluorescence (SIF), may provide more sensitive indicators of vegetation stress response to drought conditions for constructing robust joint distributions. Future work could also optimise probability criteria independently for different VD categories and regions, moving beyond a uniform conditional probability criterion (e.g., applying $P_{decision}$=40% only). Such optimisation might enhance prediction capabilities for the severe VD categories in heterogeneous environments.

Ultimately, while my framework demonstrates promise for operational drought impact prediction, particularly for no risk to moderate vegetation stress conditions, future developments should focus on enhancing extreme event prediction capabilities by understanding the mechanism of vegetation response behaviours to drought stress. The identified limitations and challenges provide clear directions for improving the robustness and applicability of impact-based drought early warning systems in southeastern Australia.

Some of these limitations are further addressed in subsequent chapters. In Chapter 4, a threshold event-matching framework is developed to enhance the prediction of low-frequency but high-severity vegetation deterioration events. Building on this, Chapter 5 integrates the copula-based probabilistic model with the event-matching approach, allowing for a more comprehensive prediction framework tailored to diverse land use types.

3.7. Conclusion

This study introduces a probabilistic framework to predict categorical vegetation deterioration risks from root-zone soil moisture deficits in southeastern Australia. I combine time-lagged correlation analysis with copula-based joint probability modelling to establish critical soil moisture trigger thresholds for an impact-based drought early warning system. These thresholds enable prediction of vegetation stress across multiple severity levels, from mild to exceptional deterioration. The following conclusions are drawn from this study,

(a) The temporal response patterns between soil moisture deficits and vegetation deterioration demonstrate strong regional heterogeneity, with response times typically ranging from 8 to 40 days. Agricultural regions, particularly dryland cropping areas, exhibited longer response times exceeding one month, while natural vegetation systems showed more rapid responses to soil moisture stress.

(b) The copula-based joint probability analysis effectively captured the non-linear relationships between soil moisture anomalies and vegetation responses. The framework successfully quantified conditional probabilities of vegetation deterioration under varying soil moisture drought scenarios, revealing that even under severe soil moisture deficits (SMA $\leqslant$ -2.5 SD), the region-wide probability of experiencing mild vegetation deterioration averaged only 61.2%.

(c) The evaluation of multiple conditional probability criteria (40% to 70%) for determining trigger thresholds revealed better performance at the 40% criterion level. The framework demonstrated strong predictive capability for no-risk to moderate vegetation stress conditions, achieving hit rates of 0.87 and 0.33 for these categories respectively. However, prediction skill diminished markedly for severe and extreme vegetation deterioration events, highlighting the challenges in forecasting low-frequency, high-impact events.

(d) The operational implementation of this framework offers several potential advantages for anticipating drought-related vegetation impacts at short temporal scale. Its probabilistic structure allows flexible adjustment of threshold criteria across different spatial domains and timeframes, which is particularly valuable in regions where soil moisture-vegetation relationships are heterogeneous. However, the reliance on joint probability modelling introduces sources of cumulative uncertainty, especially at finer temporal resolutions, where predictive skill may be compromised by limited sampling of extreme events and propagation of threshold estimation errors.

This research advances our understanding of soil moisture drought-vegetation interactions and provides a quantitative basis for impact-based drought early warning. While this study applies copula models primarily for threshold identification, the fitted dependence structures also contain information about ecosystem response behaviour, which may offer deeper ecological insights. The framework offers valuable probabilistic insights for improving agricultural and environmental management strategies in drought-prone regions. Future developments should focus on enhancing the predictive capability for low-frequent but high-impact events.

Supplementary Information for Chapter 3:

SI Tables

Table S3.1. Marginal distribution types selected for fitting soil moisture anomaly (SMA) and vegetation index anomaly (VIA) across the study region.

Distribution types	Cumulative distribution function	Parameters
Gaussian	$F(x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)dx$	μ: $(-\infty,+\infty)$ σ^2: $(0,+\infty)$
Gaussian Kernel	$F(x)=\frac{1}{2}\left[1+\mathrm{erf}\left(\frac{x-\mu}{\sqrt{2\sigma^2}}\right)\right]$	μ: $(-\infty,+\infty)$ σ^2: $(0,+\infty)$
Gamma	$F(x)=\frac{\beta^{-\alpha}}{\Gamma(\alpha)}\int_{0}^{x}t^{\alpha-1}e^{-t/\beta}dt$	α: $(0,+\infty)$ β: $(0,+\infty)$
Beta	$\int_{0}^{x}\frac{t^{\alpha-1}(1-t)^{\beta-1}}{B(\alpha,\beta)}dt$	α: $(0,+\infty)$ β: $(0,+\infty)$

Table S3.2. Mathematical formulations of the copula functions considered in the study.

Copula type	Joint Cumulative Distribution Function $C\ [F_X(x),F_Y(y)]=C[u,v]$	Generator	Parameter range θ
Gaussian	$\varphi_{\theta,\ \rho}\left(\varphi^{-1}(u),\varphi^{-1}(v)\right)$	/	(-1,1)
Clayton	$\left(u^{-\theta}+v^{-\theta}-1\right)^{-1/\theta}$	$\left(t^{-\theta}-1\right)/\theta$	$(0,+\infty)$
Frank	$-1/\theta\left[\ln\left(1+\left(e^{-\theta u}-1\right)\left(e^{-\theta v}-1\right)\right)-\ln\left(e^{-\theta}-1\right)\right]$	$\ln\left(e^{\theta t}-1\right)-\ln\left(e^{\theta}-1\right)$	$(-\infty,+\infty)\setminus\{0\}$
Gumbel	$\exp\left[-\left((-\ln u)^{\theta}+(-\ln v)^{\theta}\right)^{1/\theta}\right]$	$(-\ln t)^{\theta}$	$[1,+\infty]$

Table S3.3. Summary statistics of SMA trigger thresholds for five levels of vegetation deterioration (VD_1 to VD_5) under four exceedance probability criteria (Copula-cp40 to Copula-cp70). Values reported include the median, standard deviation, interquartile range (IQR), and sample size (number of pixels with identified thresholds)

Name of SMA sets	VD category	SMA value (SD)			
		Median	Std Dev	IQR [25%, 75%]	Sample Size
Copula-cp40	VD_1	-0.77	0.49	[-0.90, -0.50]	15368
	VD_2	-1.46	0.46	[-1.70, -1.10]	11844
	VD_3	-2.08	0.42	[-2.40, -1.70]	5242
	VD_4	-2.3	0.38	[-2.60, -2.00]	3201
	VD_5	-2.48	0.32	[-2.70, -2.20]	2509
Copula-cp50	VD_1	-1.11	0.53	[-1.30, -0.70]	13962
	VD_2	-1.7	0.46	[-1.90, -1.40]	9233
	VD_3	-2.2	0.44	[-2.60, -1.80]	4014
	VD_4	-2.35	0.37	[-2.70, -2.00]	2761
	VD_5	-2.51	0.31	[-2.80, -2.30]	2239
Copula-cp60	VD_1	-1.36	0.54	[-1.60, -0.90]	11759
	VD_2	-1.93	0.45	[-2.20, -1.60]	6629
	VD_3	-2.26	0.44	[-2.60, -1.90]	3302
	VD_4	-2.39	0.36	[-2.70, -2.10]	2431
	VD_5	-2.53	0.29	[-2.80, -2.30]	1976
Copula-cp70	VD_1	-1.58	0.52	[-1.90, -1.20]	9341
	VD_2	-2.13	0.45	[-2.50, -1.80]	4449
	VD_3	-2.3	0.42	[-2.70, -1.90]	2668
	VD_4	-2.42	0.34	[-2.70, -2.10]	2144
	VD_5	-2.56	0.28	[-2.80, -2.40]	1774

SI figures

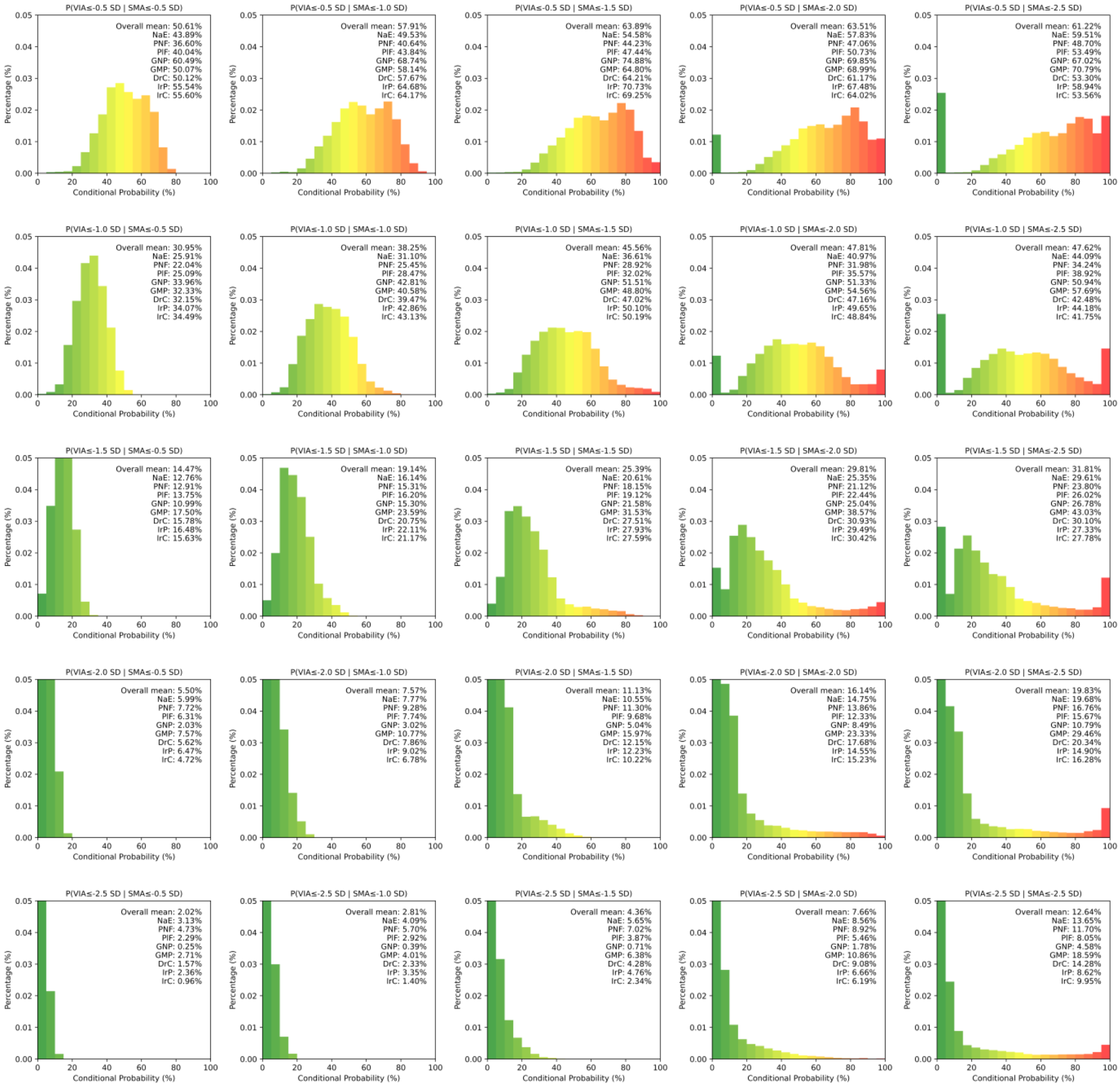


Fig. S3.1. Summary statistics of conditional probabilities of vegetation deterioration severity levels (VIA ≤ z-score thresholds) under different soil moisture drought scenarios (SMA ≤ z-score thresholds), averaged across all pixels. Bars represent the mean conditional probabilities at each VIA threshold level under mild to exceptional SMA categories.

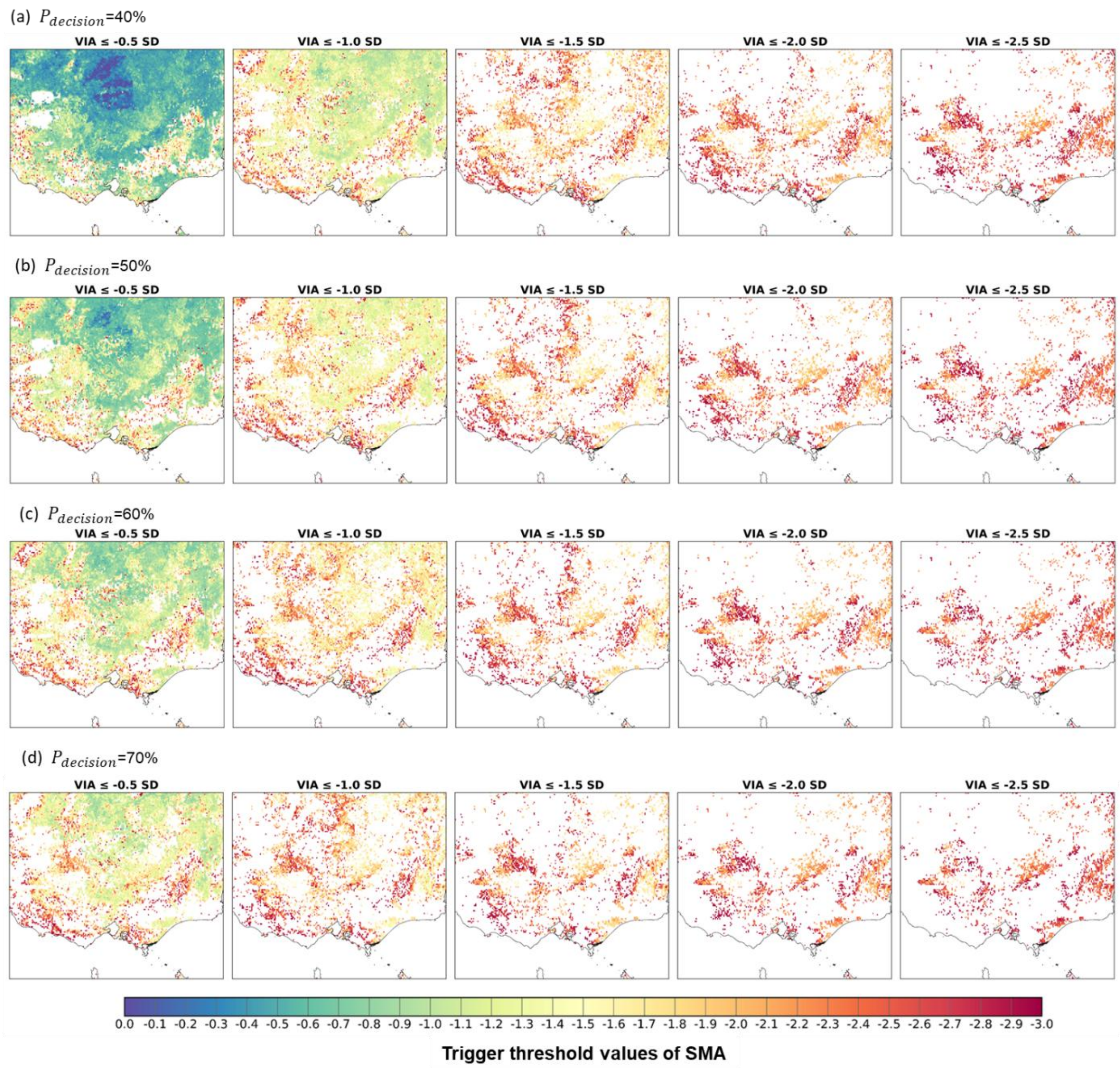


Fig. S3.2. Spatial distribution of soil moisture anomaly (SMA) trigger thresholds for five vegetation deterioration (VD) severity levels (VIA ≤ -0.5 to -2.5 SD) under different exceedance probability criteria: (a) 40%, (b) 50%, (c) 60%, and (d) 70%. Each panel shows the minimum SMA deficit required to trigger a given level of VD with the specified probability. Warmer colours (e.g., red to purple) indicate more severe soil moisture deficits (i.e., lower SMA values). White areas indicate regions where no valid threshold was identified.

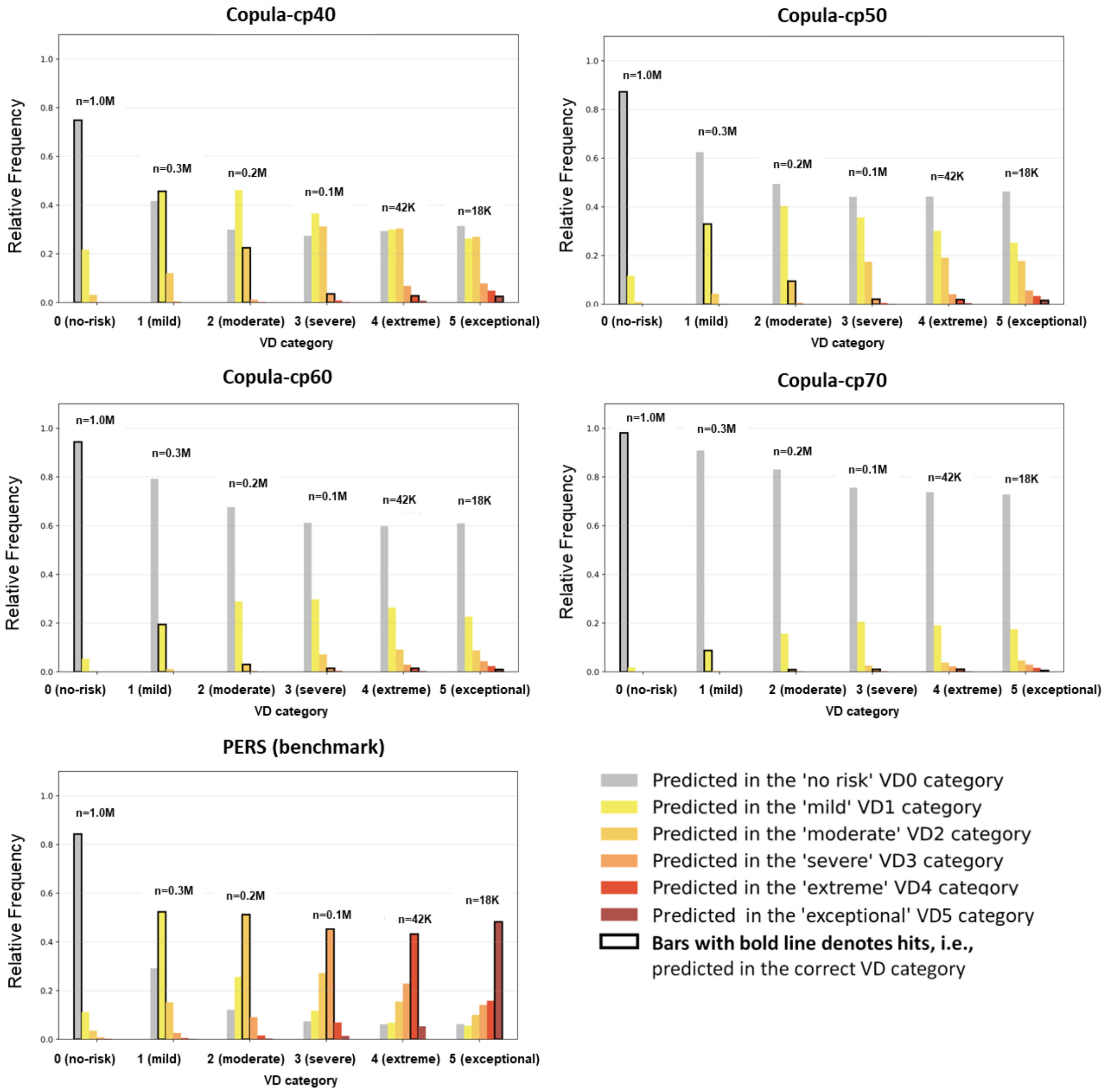


Fig. S3.3. Same as Fig. 3.12 but for monthly time scale.

Chapter 4

Skill and lead time of vegetation drought impact predictions based on soil moisture observations

The content of this chapter is based on the following article published in *Journal of Hydrology*:

Li, Y., van Dijk, A. I., Tian, S., & Renzullo, L. J. (2023). Skill and lead time of vegetation drought impact forecasts based on soil moisture observations. Journal of Hydrology, 620, 129420. https://doi.org/10.1016/j.jhydrol.2023.129420.

Abstract

Timely and skilful forecasts of vegetation drought impacts should enable more proactive drought preparedness, management and mitigation. Numerous previous studies have found a temporal correlation between soil moisture and vegetation condition. However, a correlation across the full range of soil moisture and vegetation condition does not automatically translate into skill in forecasting infrequent events such as agricultural droughts. Here, I develop a threshold- or impact-based forecasting framework to assess early warning capability (EWC). I analyse the skill and lead time achieved using soil moisture observations at multiple depths as predictors of subsequent vegetation drought impacts inferred from MODIS satellite observations at 93 sites across the United States. Forecast thresholds are expressed in terms of seasonally-adjusted standardised anomalies (z-scores) to distinguish climate-related impacts from any seasonal vegetation cycle. Different combinations of soil moisture integration depth, satellite vegetation indices and threshold levels are tested. Near-Infrared Reflectance of vegetation (NIRv) yields a marginally better EWC than other indicators of vegetation drought impact. The optimal soil moisture integration depth varies between land cover types, from 0-10 cm for cropping systems to 0-100 cm for grassland and shrubland. The greatest skill improvements are achieved using similar z-score threshold values for the soil moisture trigger and vegetation impact, producing typical lead times of two to four weeks. Further research is recommended to combine the framework developed here with spatially continuous soil moisture analyses or forecasts available from remote sensing and land surface models.

4.1. Introduction

As emphasised in Chapters 2 and 3, drought is among the costliest natural hazards, with far-reaching impacts on ecosystems, society, and economies (Sheffield & Wood, 2012; van Dijk *et al.,* 2013). Agriculture bears the brunt of these impacts, accounting for over 80% of total drought-related economic losses (FAO, 2021). Drought frequency and severity are projected to increase under climate change, posing growing risks to global food security (IPCC, 2015). Agricultural drought typically manifests when soil moisture deficits reach levels that trigger visible stress responses in crops and pastures—such as leaf wilting, reduced chlorophyll content, and stunted growth—thereby reducing yield and forage quality (Asoka & Mishra, 2015; Flexas *et al.,* 2000; Vicente-Serrano *et al.,* 2012).

Satellite-based vegetation indices (VIs), such as NDVI, VCI, and VHI, have long been used as proxies to monitor drought impacts across large regions (Boken *et al.,* 2005; Kogan, 1995, 2002). As discussed in detail in Chapter 2, Section 2.2.2, these indices reflect vegetation stress only after it manifests, making them less effective for anticipatory drought management (Kogan *et al.,* 2015). The increasing focus on drought preparedness has therefore shifted attention from reactive impact monitoring toward early warning systems (EWS) that aim to forecast vegetation stress before it becomes visible (Coughlan de Perez *et al.,* 2015; WMO, 2015).

As discussed in Chapter 3, soil moisture anomalies, particularly in the root-zone depth (0-1m), have been shown to be more closely correlated with vegetation condition than simple meteorological indices (Bachmair *et al.,* 2016a, 2016b; Liu *et al.,* 2020). Most existing studies, however, have relied on correlation-based analyses to estimate forecast potential, using either Pearson/Spearman coefficients or anomaly correlation (e.g., Asoka & Mishra, 2015; Tian *et al.,* 2019). While useful, these measures reflect average associations and do not necessarily indicate whether extreme vegetation deterioration events can be predicted in advance. This is especially problematic because drought impacts are low-frequency events, concentrated in the lower tail of the soil moisture and VI distributions. Strong overall correlation does not imply reliable skill in predicting such rare, high-impact occurrences (Brust *et al.,* 2021; Hao *et al.,* 2017a).

Building upon the research focus established in Chapters 2 and 3, this chapter specifically addresses Research Question 2 (see Chapter 1, Section 1.3):

Do in-situ soil moisture observations, through the threshold event-matching approach, offer potential capability in predicting infrequent-severe vegetation deterioration events, and if so, what is the lead time?

While previous chapters explored the statistical association between soil moisture anomalies and vegetation stress, particularly using copula-based modelling to estimate multi-categorical deterioration risks. This chapter shifts focus toward predictive ability of rare and severe events. This directly contributes to the second thesis-wide objective: *developing a process-informed predictive framework that supports operational drought early warning.*

To that end, I propose a *threshold-based event-matching framework* that extends beyond prior correlation-based approaches by providing categorical skill evaluation across drought severity levels. As predictors, I use in situ soil moisture measurements at different depths from 93 stations across the USA. For predictands, I assess a set of optical vegetation indices derived from MODIS observations. Both time series are de-seasonalised and standardised into z-scores to enable generalisation across sites and scales and further extended application of predictive framework. Predict performance is evaluated using the F_2-score, which prioritises reduced missed detections while tolerating moderate false alarms. This chapter aims to address the following specific research questions:

(1) Do soil moisture observations have skill in forecasting subsequent vegetation drought impacts, and if so, what is the lead time?
(2) What combinations of standardised soil moisture and vegetation anomalies – 'z-score threshold combinations' hereafter – can forecast vegetation drought impacts of increasing severity?
(3) How do soil moisture integration depth, vegetation type and climate regime affect skill and lead time?

4.2. Data and methodology

4.2.1. In situ soil moisture observations

In situ soil moisture data for 1 January 2003 to 31 December 2019 were collected from three national and two regional networks across the continental United States of America (Table 4.1). The Soil Climate Analysis Network (SCAN; https://www.wcc.nrcs.usda.gov/scan/) and the United States Climate Reference Network (USCRN; https://www.ncdc.noaa.gov/crn/) provide hourly neutron probe measurements of volumetric soil water content at five standard depths (5, 10, 20, 50, and 100 cm). The SCAN sites are mostly located in agricultural regions (Schaefer *et al.,* 2007), whereas the USCRN sites are associated with climate monitoring stations and likely to retain consistent land cover type over decades (Bell *et al.,* 2013). The AmeriFlux sites (https://ameriflux.lbl.gov/sites/) are located at eddy covariance flux towers that measure carbon and water exchanges (Novick *et al.,* 2018). Only a few flux sites were equipped with volumetric soil water content sensors that monitored half-hourly at various soil depths (typically 5, 10, 30 and 50 cm). The Illinois Climate Network (ICN; https://www.isws.illinois.edu/warm/soil/) was established to support the agriculture sector by providing volumetric soil water content at four depths (5, 10, 20 and 50 cm) (WARM 2020). The Oklahoma Mesonet (OKM; https://www.mesonet.org) is an environmental monitoring network. The fraction of soil saturation was measured at four depths (5, 25, 60 and 75 cm) in OKM based on heat capacity measurements at sub-hourly time steps (e.g., 5 and 30 minutes) (Brock *et al.,* 1995; McPherson *et al.,* 2007).

Table 4.1 Details of in situ soil moisture data used in the study

In situ network	Study regions	Record year	Measurement unit	Sensor depths (cm)	Reference
Soil Climate Analysis Network (SCAN)	contiguous USA	1991-present	volumetric soil water content	5, 10, 20, 50, 100	Schaefer et al. (2007)
Unites States Climate Reference Network (USCRN)	contiguous USA	2008-present	volumetric soil water content	5, 10, 20, 50, 100	Bell et al. (2013)
Ameriflux	contiguous USA	1997-present	volumetric soil water content	5, 10, 30, 50	Novick et al. (2018)
Illinois Climate Network (ICN)	Illinois, USA	1989-present	volumetric soil water content	5, 10, 20, 50	WARM (2020)
Oklahoma Mesonet (OKM)	Oklahoma, USA	1994-present	fraction of saturated soil	5, 25, 60, 75	Brock et al. (1995); McPherson et al. (2007)

Sites were selected from each network based on the following criteria: (1) record continuity and length, (2) spatial homogeneity, (3) the presence of rainfed (non-irrigated) vegetation, and (4) measurement depth. First, I calculated the daily-average soil moisture at each sensor depth using the hourly or sub-hourly observations with the highest quality flag. Missing daily averages were estimated using linear interpolation, but records missing more than 15% of the daily averages in each year were excluded. I only considered sites that had records for at least six years. Second, at each site, spatial homogeneity within a radius of 500 m was investigated using high-resolution imagery (Google Earth Pro version 7.1) and the MODIS International Geosphere-Biosphere Programme (IGBP) land cover product (MCD12Q1) (Friedl and Sulla-Menashe, 2019). I only retained sites with a stable and homogeneous land cover that had not been disturbed by fire occurrences during the study period. Third, I excluded sites on cultivated land equipped with a centre-pivot irrigation system or land cover with significant bare soil patches (e.g., open shrubland). Fourth, only sites with observations available to a depth of at least 50 cm were retained.

A total of 93 sites satisfied all criteria (Fig. 4.1): 33 OKM sites, 25 USCRN sites, 20 SCAN sites, nine ICN sites and six Ameriflux sites. Table S4.1 (see Supplementary Information) lists the geographic coordinates, land-cover types (MODIS IGBP), aridity index, climate regime, and temporal coverage of the 93 sites. They include 30 sites with 6 to 10 years of data and 63 sites with 11 to 17 years. Based on the IGBP land cover product, the 93 sites were categorised into five ecosystems: grasslands (GRA, 38 sites with 506 site-years), croplands (CRO, 16 sites with 226 site-years), cropland/ natural vegetation mosaics (CVM, 17 sites with 225 site-yeas), savannas (SAV, 11 sites with 151 site-years), and woody savannas (WSA, 11 sites with 126 site-years). Several sites were located in natural ecosystems without any agricultural activity. I retained these for comparative purposes.

For each site, the soil moisture observations at multiple measurement depths were weighted to compute depth-integrated soil moisture (θ_d) for the upper root zone (d=10 cm), shallow root zone (d=30 cm), and deep root zone (d=100 cm) at a daily time step. The integration depths of soil moisture were loosely based on the vertical root distribution in different ecosystems (Arora and Boer, 2003), as well as soil layer definitions commonly used in global products, such as GLDAS Noah (0-10, 10-40, and 40-100 cm) (Rodell *et al.,* 2004) and ERA5-Land (0-7, 7-28, and 28-100 cm) (Muñoz-Sabater *et al.,* 2021). The weights were determined by assuming that each sensor measured soil moisture for a layer bound by half the distance between the next higher and lower sensors, where applicable.

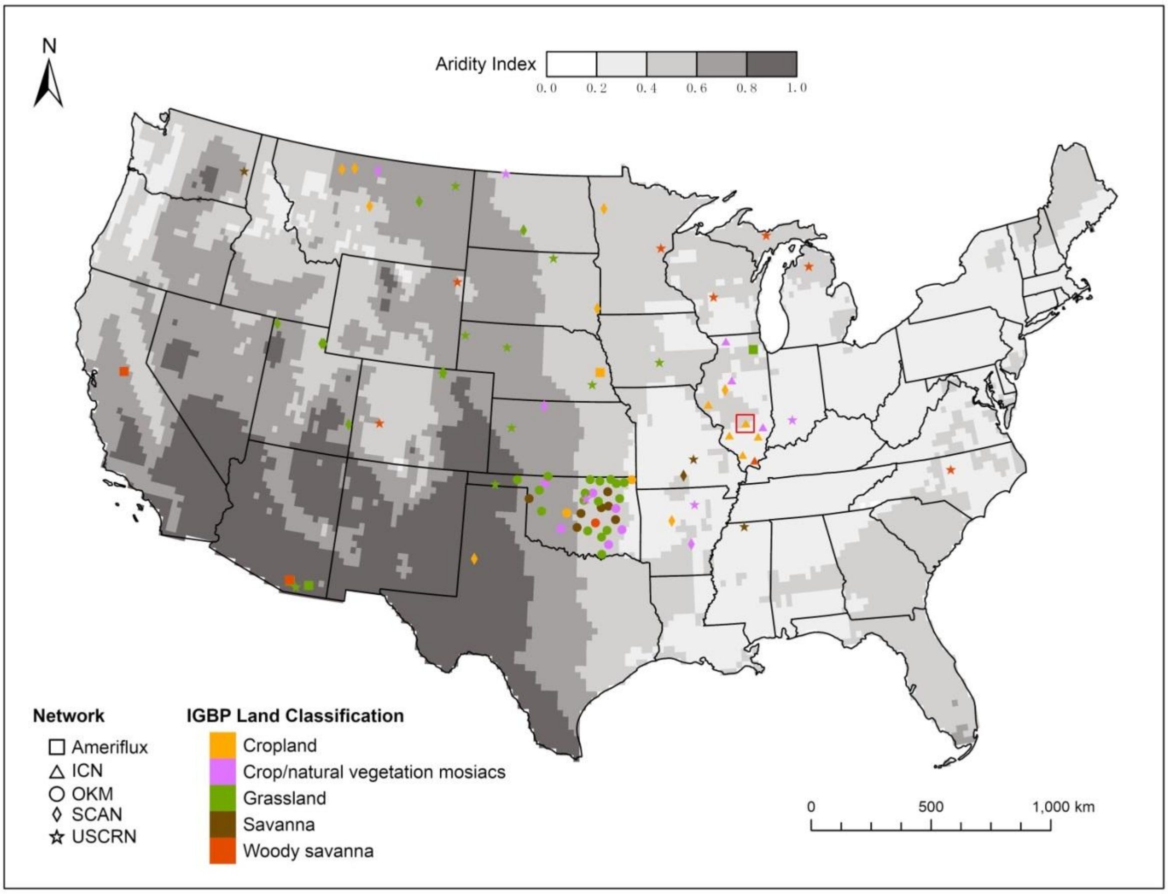


Fig. 4.1. Locations of the 93 in situ soil moisture observation sites across the continental US. The symbol shapes indicate the networks, and the filled colours indicate the MODIS IGBP land-cover classifications. Base map shows the Aridity Index over the reference period 1976-2000. The red square maker denotes location for which time series will be further shown in detail below in Fig 4.2.

4.2.2. Satellite-derived vegetation indices

I evaluated four satellite-derived vegetation indices (VIs): the Normalised Difference Vegetation Index (NDVI, Tucker, 1979), the Enhanced Vegetation Index (EVI, Huete *et al.,* 2002), the Normalized Difference Infrared Index (NDII, Yilmaz *et al.,* 2008), and the Near-Infrared Reflectance of vegetation (NIRv, Badgley *et al.,* 2017). NDVI and EVI are two of the most widely used proxies of vegetation growth, condition or vigour (Xue & Su, 2017). NDII has been used to estimate the equivalent water thickness of leaves and canopy (Yilmaz *et al.,* 2008). The NIRv index was recently proposed as a better predictor of vegetation productivity (Badgley *et al.,* 2017; Baldocchi, 2020). All VIs can be calculated using red, blue, near-infrared and shortwave-infrared reflectance measurements from the MODIS sensors (Table

4.2). I extracted atmospherically-corrected nadir BRDF-Adjusted Reflectance (NBAR) for each site from the Collection 6 Terra and Aqua MODIS MCD43A4 daily data set for the 500-m pixel centred geographically closest to the reported soil moisture sites. The daily MODIS data represent a weighted average of daily measurements over 16 days, with the ninth day as the reporting date (Schaaf & Wang, 2015). Hence reflectances can be influenced by vegetation condition before and after the nominal date, depending on atmospheric conditions on different days within the period. My study included 17 years (2003-2019) for which both Terra and Aqua MODIS data were available. I only used retrievals marked as “highest quality” in all bands (i.e., Bitmask = 0 in the quality assurance grid). I excluded periods when NDVI was less than 0.15 to reduce the effect of snow, ice and bare soil.

Table 4.2 MODIS vegetation indices tested in this study, where B1, B2, B3, and B6 are surface reflectance in MODIS band 1 (620-670 nm), band 2 (841-876 nm), band 3 (459-479 nm) and band 6 (1628-1652 nm), respectively.

Vegetation index	Acronym	Formula	References
Normalised Difference Vegetation Index	NDVI	NDVI = (B2-B1)/(B2+B1)	Tucker (1979)
Enhanced Vegetation Index	EVI	EVI = 2.5(B2-B1)/(B2+6B1-7.5B3+1)	Huete *et al.,* (2002)
Normalised Difference Infrared Index	NDII	NDII = (B2-B6)/(B2+B6)	Yilmaz *et al.,* (2008)
Near Infrared Reflectance of vegetation	NIRv	NIRv = B2(B2-B1)/(B2+B1)	Badgley *et al.,* (2017)

4.2.3. Definition of threshold soil moisture and vegetation condition events

I wished to remove any recurrent seasonal cycle that might artificially inflate skill estimates. First, the mean of the four VIs and three θ_d (i.e., θ_{10}, θ_{30}, and θ_{100}) series was calculated for each day of the year (DOY, *t*) for their common period at each site. This average seasonal pattern was adopted as the estimated average climatological cycle or climatology. The anomalies of VI and θ_d were subsequently calculated as the difference between the observed values and climatology for the corresponding day *t*:

$$a_x(t,y) = x(t,y) - \bar{x}(t) \tag{4.1}$$

where $\bar{x}(t)$ is the period-average value for day *t*, $x(t,y)$ the daily value of the variable for day *t* of year *y*, and $a_x(t,y)$ the anomaly for day *t* of year *y.*

The anomalies of VI and θ_d were assumed to follow a normal distribution. To make it easier to compare anomalies in various variables, I standardised the anomalies of variables by subtracting the period mean anomaly at each time step t and then dividing by the period standard deviation of the anomaly. The generic calculation of z-scores for each variable was:

$$z_x(t) = a_x(t, y)/\sigma_{ax}(t) \tag{4.2}$$

where $a_x(t, y)$ is the anomaly for day t of year y, $\sigma_{ax}(t)$ the standard deviation of anomalies for day t, and $z_x(t)$ the standardised anomaly or z-score for day t.

The z-score indicates how many standard deviations an anomaly is away from the mean seasonal cycle. A negative score implies that the deviation is below the average. Thus, negative z-scores for soil moisture anomalies, $z(\theta)$, represent drier-than-average soil moisture conditions, and negative z-scores for VI anomalies, $z(VI)$, represent a decline in vegetation condition. Fig. 4.2 shows an example of daily time series of z-scores for NIRv versus three θ_d depths at a cropland site in the mid-central United States.

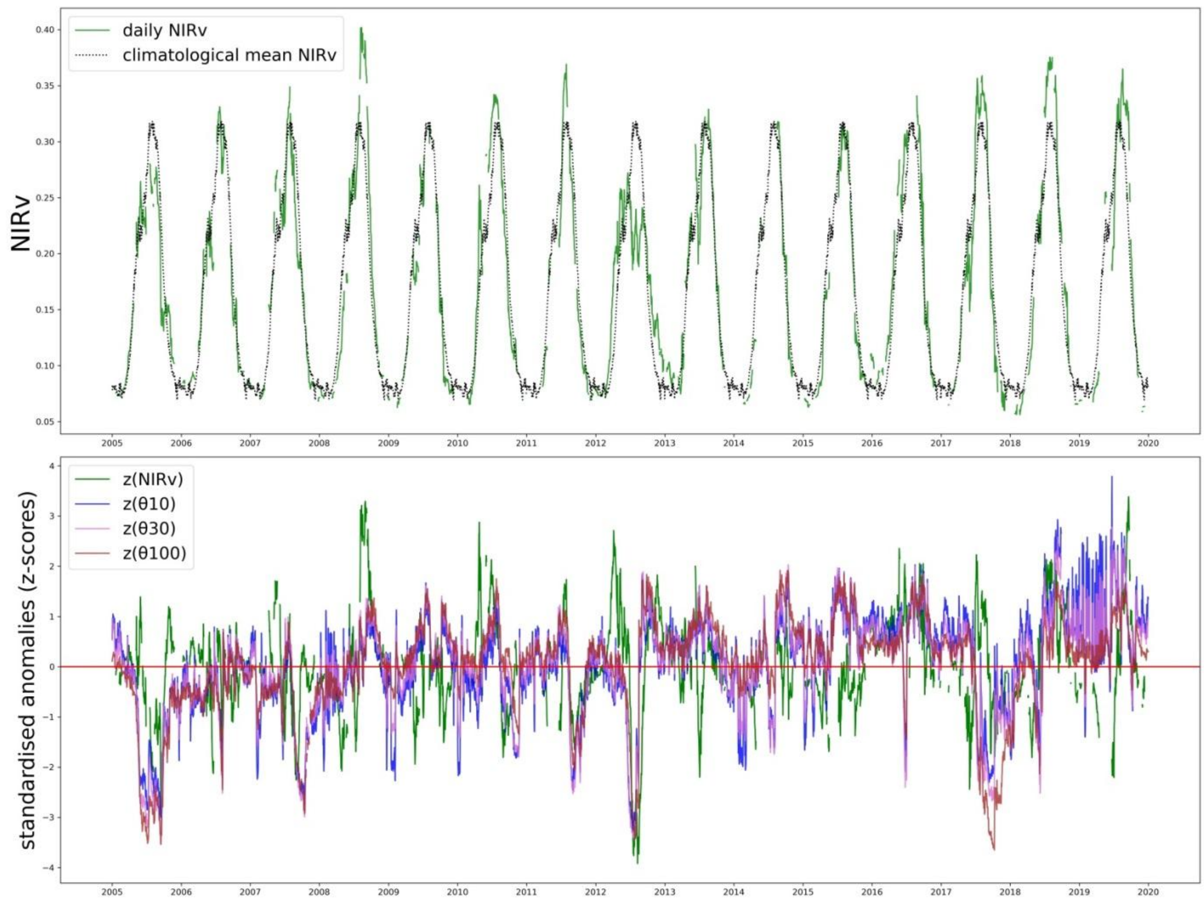


Fig. 4.2. Time series plots of daily NIRv (top in green), climatological mean NIRv (top in dashed black), and the seasonally-adjusted and standardised anomalies (z-scores) from 2005 to 2020 for NIRv (bottom in green) against the three depth-integrated soil moisture for the upper root zone (θ_{10}; bottom in blue), shallow root zone (θ_{30}; bottom in purple), and deep root zone (θ_{100}; bottom in brown) for an example site of the Illinois Climate Network (ICN) Brownstown (longitude = -88.9, latitude=38.9) shown in red box of Figure 4.1.

I established a framework to identify θ_d falling below a defined threshold using $z(\theta)$ and the onset of vegetation drought impacts using $z(VI)$. I subsequently calculated two key performance metrics, *Lead time* (Δt) and the overall skill score (F_2-score), to evaluate the early warning capability (EWC). I restricted the analysis to the peak growing season to avoid the influence of phenological or temperature-driven vegetation browning and senescence. The peak growing season was determined by the land surface phenology approach, where phenological dates were determined using dynamic thresholds or a particular derivative alongside the climatological curve of the VI (Zeng *et al.,* 2020). The mid-green-up and mid-brown-down points were the day of the year (DOY) at which the smoothed VI climatology crosses above and again below 50% of the VI amplitude, respectively (Fig. 4.3), and the peak growing season was defined as the period between mid-green-up and mid-brown-down.

The pre-green date was defined as the DOY with the trough of VI climatology closest to the green-up period. If there were two phenological periods at a given site, I only considered the period with the larger amplitude.

I defined the onset of vegetation drought impact as the first date (DOY) during the peak growing season on which $z(VI)$ fell below a chosen z threshold. The soil moisture deficit should precede the vegetation impact. Therefore, I iteratively extended the search window backwards until the pre-green date and determined the onset of soil moisture deficiency as the last day preceding the impact event on which $z(\theta)$ fell below the specified threshold. Combining the predictor and predictand *z-scores*, I analysed a total of 36 *z* threshold combinations [$z(\theta)$, $z(VI)$], encompassing all *z* combinations between -2.5 and 0.0 at an increment step of 0.5. The experiment was repeated for all twelve combinations of the four VIs and three θ depths.

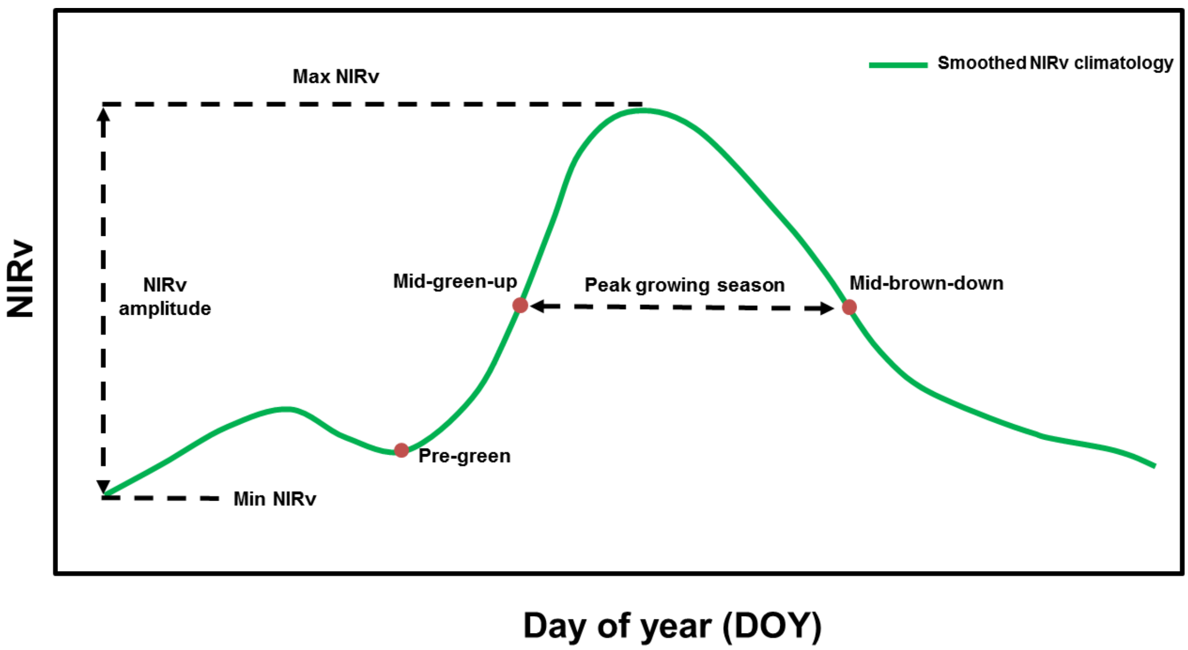


Fig. 4.3. Illustration of identification of climatological transition points (mid-green-up and mid-brown-down) using 50% of NIRv amplitude as the threshold value. The DOY nearest to the green-up period is the pre-green point.

4.2.4. Evaluation of early warning capability

I defined the first occurrence of VI threshold exceedance of each site-year as the onset of drought vegetation impact (Fig. 4.4). If the $z(\theta)$ threshold was crossed before the $z(VI)$ threshold (Fig. 4.4a), it was deemed a true positive (*TP*) occurrence that successfully triggered an early warning. I found that $z(\theta)$ could fall below the threshold multiple times within the search window before a VI impact. In such cases, I only regarded the last preceding $z(\theta)$ threshold occurrence as a successful early warning, while earlier occurrences were considered false positives (*FP*), triggering false alarms. If $z(\theta)$ did not cross the threshold at all before vegetation decline (Fig. 4.4b), the event was classified as a false negative (*FN*). Most of the time series is made up of days without drought conditions. They can be considered true negatives (*TN*), but do not give a meaningful assessment of skill. Hence a skill metric that does not include *TN*, the F_2-score, was chosen to quantify skill. For each site-year, the true positives (*TP*), false positives (*FP*), and false negatives (*FN*) were counted. The F_2-score requires estimates of *Precision* and *Recall*, defined as:

$$Precision = TP / (TP + FP) \tag{4.3}$$

$$Recall = TP / (TP + FN) \tag{4.4}$$

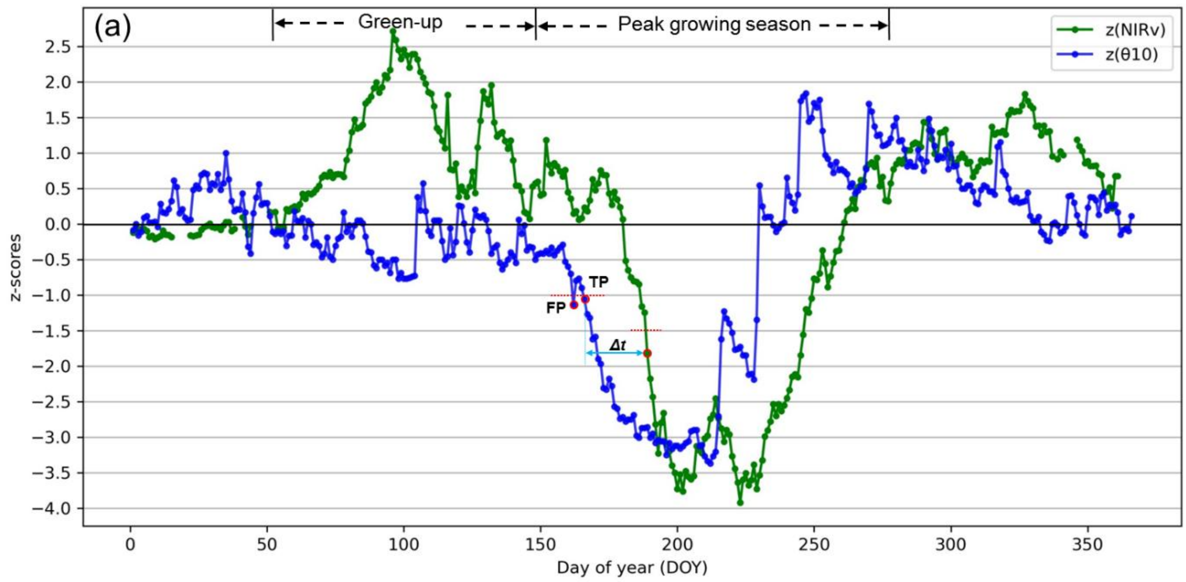
where *TP*, *FP* and *FN* are the number of true positives, false positives, and false negatives, respectively. *Precision* describes the likelihood that a $z(\theta)$ event accurately anticipates a $z(VI)$ event and penalises the measures for false alarms (*FP*). *Recall* describes the proportion of vegetation decline events that were accurately anticipated by a $z(\theta)$ trigger and penalises missed $z(VI)$ events (*FN*). The following skill score captures the trade-off between precision and recall:

$$F_\beta\text{-}score = (1+\beta^2) \cdot (Precision \cdot Recall) / (\beta^2 \cdot Precision + Recall) \tag{4.5}$$

The value of β is subjective and is chosen to reflect that *Recall* is considered β times as important as *Precision*. Values of 0.5, 1, and 2 are commonly used for β. I preferred overpredicting vegetation impact events rather than missing them. I selected a value of 2, which weighs *Recall* twice as much as *Precision* to provide a composite score for the skill metric.

The resultant F_2-score ranges from 0 to 1, with a higher score indicating better forecast skill. The F_2-score is a relative skill score with no specified standards for comparing it to a normative quality assessment (e.g., good, acceptable, or poor). However, it can be used to compare the performance of different predictors and predictands within the same forecasting approach, as done here. For each z threshold combination, *Precision*, *Recall* and F_2-scores are calculated separately for the three θ depths and four VIs. The results were compared and stratified according to land cover type and climate regime. For each z threshold combination, the *Lead time* (Δt) between the $z(\theta)$ trigger and the $z(VI)$ event was computed for each site and year. These four predictive performance metrics (*Precision, Recall,* F_2-score and *Lead time*) provide complementary insights into how effectively the $z(\theta)$ thresholds identify $z(VI)$ events across different standardised severity levels.

Due to the nature of the forecasting framework and the formulation of the F_2-score, there is no direct method to determine whether a calculated skill score is significantly different from a naïve forecast. Instead, I performed a control forecast experiment for which I expected no skill on theoretical grounds and could calculate a baseline or background skill score. For the control experiment, I sought to remove any temporal relationship between predictor and predictand by switching the first and the second part of the $z(\theta)$ time series without doing the same for $z(VI)$. The first part was extended by half a year in the case of an odd number of years. I consider the skill of the control forecast equivalent to that of a random forecast, at least on average across several sites. The average change in skill score between the original and control experiment can therefore be considered as the improvement from an unskilled forecast. A statistical t-test was applied for each experiment to determine whether the average change among sites ($N=93$) was statistically significantly different from zero at $p<0.05$.



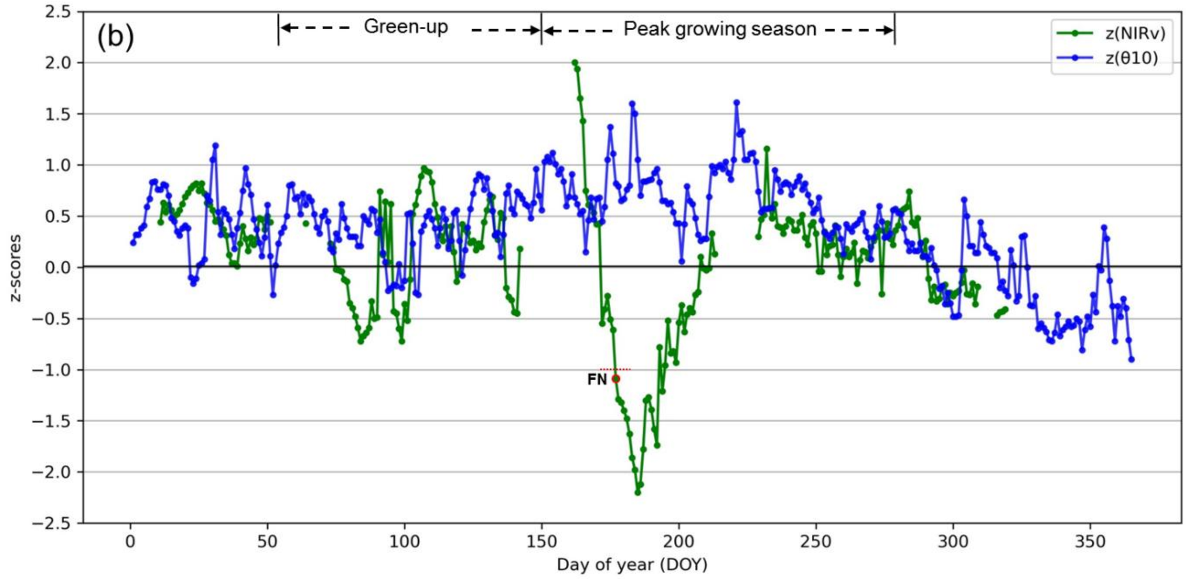


Fig.4.4. The schematic example of identifying the number of true positives (*TP*), false positives (*FP*), false negatives (*FN*), and *Lead time* (Δt) using specific zscore-threshold value for the soil moisture deficiency triggers and vegetation drought impacts at ICN-Brownstown site in (a) year 2012 and (b) year 2013. When applying the threshold of -1.0 to $z(\theta_{10})$ as the trigger in warning the threshold of -1.5 to *z(NIRv)* in year 2012, there was one false positive (*FP*) warning and one true positive (*TP*) warning, and the *Lead time* (Δt) was calculated as the difference between the onset day of year (DOY) when $z(\theta_{10})$ and *z(NIRv)* crossing the applied threshold. In year 2013, the $z(\theta_{10})$ failed in early warning the vegetation drought impacts represented by *z(NIRv)*, in this case, there was neither *TP* warning nor *FP* warning, but one false negative (*FN*) counted.

4.2.5. Analysis between vegetation types and climate regimes

Access to deeper soil moisture through greater rooting depth and the degree of adaptation to drought stress vary between vegetation types. We used the earlier mentioned IGBP land cover product to evaluate whether there were significant differences in aspects of forecast skill between different vegetation types. I expected the shallowest rooting depth and the greatest vulnerability to dryness for croplands (CRO, N=16) and cropland/natural vegetation mosaics (CVM, N=17). Conversely, I anticipated woody savannas (WSA, N=11) might have the greatest rooting depth and be the least sensitive to dry conditions, with grasslands (GRA, N=38) and savannas (SAV, N=11) having intermediate characteristics.

Climate regime, in particular aridity, may be a key factor in determining viable vegetation growth form, but also shapes other, less pronounced physiological adaptations to water availability. Therefore, I also compared the performance metrics of early warning capability for different aridity regimes. The Aridity Index was calculated as the ratio of mean precipitation to mean potential evapotranspiration (PET), estimated using the FAO-56 Penman–Monteith method (Allen et al., 1998), and classified dryness levels following Tian et al. (2019).

4.3. Results

4.3.1. Temporal characterisation of threshold events

Some of the characteristics of threshold events identified at different $z(\theta)$ and $z(VI)$ threshold levels are shown in Fig. 4.5. Depending on the thresholds chosen, events occurred between 0.25 to 8 times per year (Fig. 4.5a), corresponding with return intervals of ~0.1 to 4 years. For a chosen threshold level, $z(\theta)$ events tended to occur more frequently than $z(VI)$ events, especially for θ_{10}. This can be explained by the greater short-term variability in this shallow layer compared to greater soil depths or the larger-scale and temporally-composited VIs. These patterns were consistent between the four VIs. For the $z(\theta)<-2.5$ and $z(VI)<-2.5$ thresholds, the respective events showed a similar frequency of occurrence.

The number of sites reaching different thresholds for individual years during the 2003-2019 period shows a pattern of relatively wetter and dryer conditions across the 93 sites. Similar patterns were observed for $z(\theta)$ and $z(VI)$ (Fig.4.5b and 4.5c). The number of dryness-affected sites was largest in 2012 (80%), while relatively high numbers were also observed in 2006 (65%) and 2011 (50%). Near normal and moderate threshold events ($z(\theta)$<-1.5 to $z(\theta)$ < 0.0) occurred at the majority of the sites (>70%) in nearly every year. Only one in ten sites were affected by severe events ($z(\theta)$<-2.5 and $z(VI)$<-2.5) during most years.

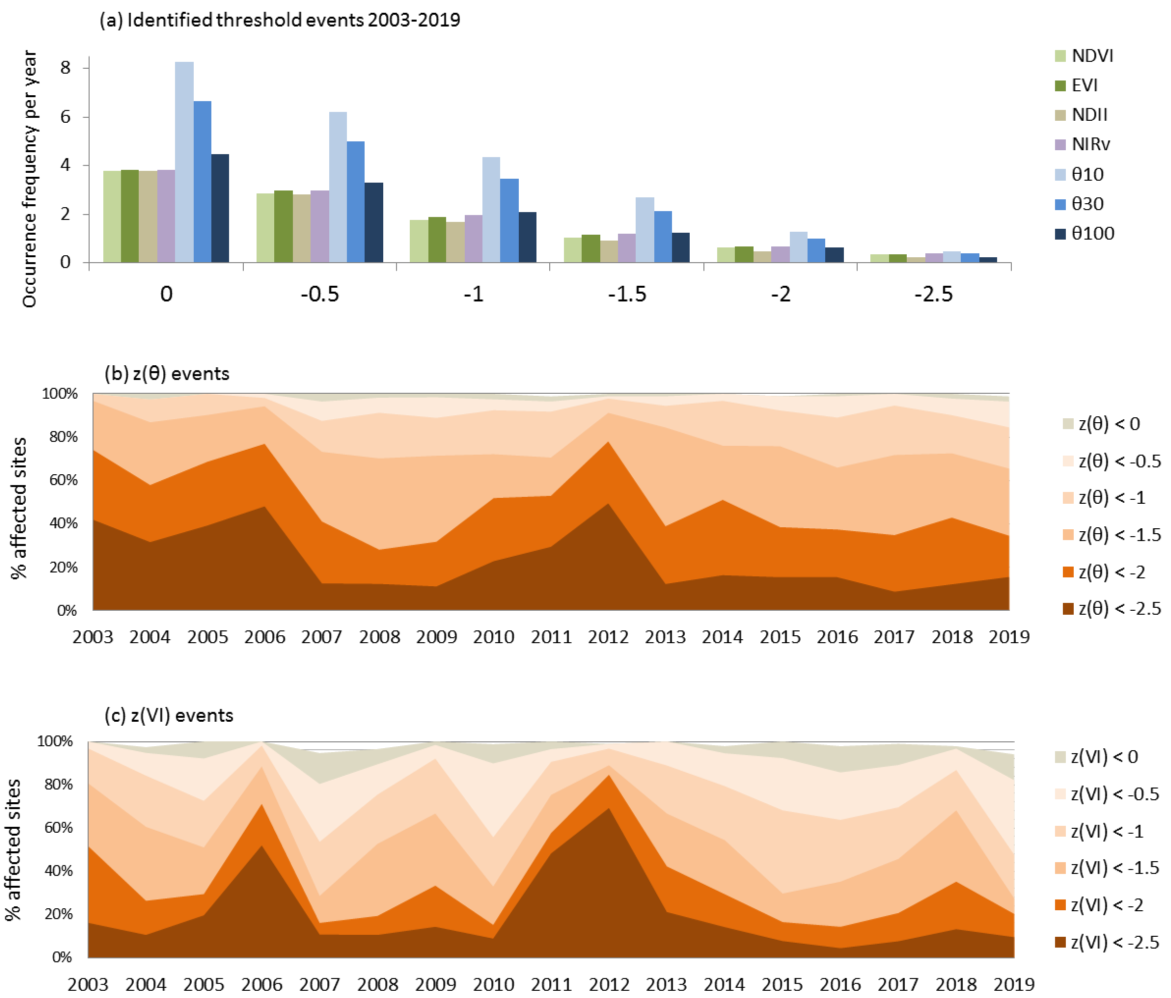


Fig.4.5. Characteristics of threshold events identified at different *z(θ)* and *z(VI)* threshold levels from 0 to -2.5 SD: (a) occurrence frequency across the selected sites; (b and c) percentage of in situ sites affected by varying *z(θ)* and *z(VI)* threshold events during the 2003-2019 period.

4.3.2. Evaluation of potential capability in forecasting different proxies of vegetation degradations

Fig. 4.6 presents the improved forecast skill score and expected *Lead time* for selected in situ sites. For any z threshold combination, forecast potential was indicated where the F_2-score showed a statistically significant improvement ($p < 0.05$) compared to the skill of the baseline control experiment across the 93 sites. The experiment was repeated 432 times, i.e., for three predictors (θ_{10}, θ_{30} *and* θ_{100}) times four predictands (NDVI, EVI, NDII, and NIRv) times 36 (six times six) z threshold combinations.

The distribution of F_2-scores indicates that forecasting skill was poorer when using NDII compared to the other three VIs, with only marginal differences between NDVI, EVI, and NIRv (Fig. 4.6a). The median value of expected *Lead time* was 12 to 30 days (Fig. 4.6b) without apparent differences among the four VIs. The greatest soil moisture integration depth showed both higher skill and lead time. However, 38 out of 93 (i.e., 41%) sites are located on grassland, and this dominance may have affected the result (see Section 4.3.4).

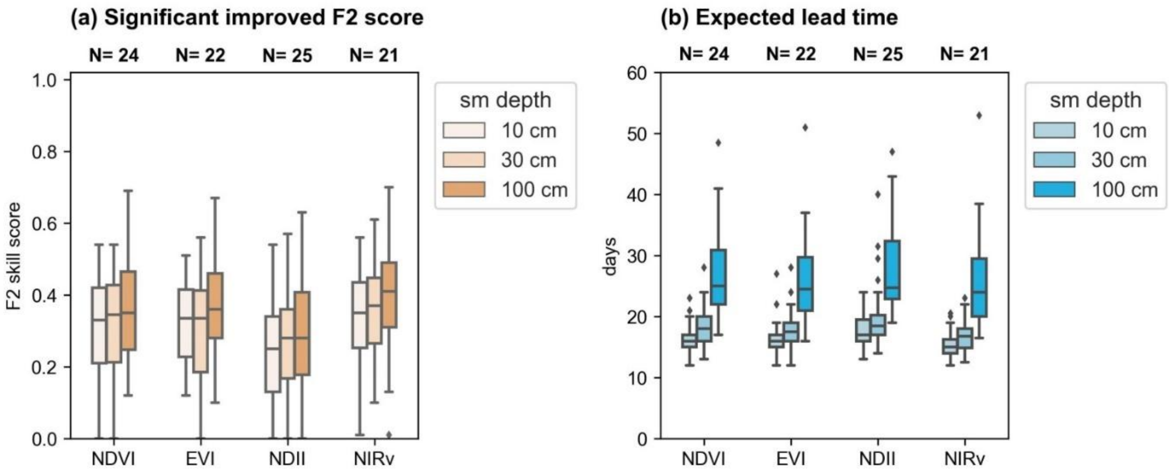


Fig. 4.6. Box and whisker plot representing the distribution of (a) significant improved F_2-score and (b) expected *Lead time* of forecasting vegetation threshold events from three integrated depths (10, 30, and 100 cm) over four vegetation indices (NDVI, EVI, NDII, and NIRv). The *N* denotes the number out of 36 z threshold combinations significantly improved from the control experiment. The outliers are marked as diamonds.

4.3.3. Influence of z threshold combinations

Different [$z(\theta)$, $z(VI)$] threshold combinations produce different skills and lead times. In presenting these differences, I focused on the results achieved with the vegetation condition index NIRv, because it showed marginally higher F_2-scores among the four VIs. The F_2-score, its components *Recall* and *Precision*, and the *Lead time* (Δt) for different z threshold combinations and three soil layer depths were calculated for all 93 sites. Subsequently, the same metrics were calculated for the control experiment, and the improvement and the statistical significance of that improvement were calculated. The numbers shown in Fig. 4.7 represent median values of the distribution (N=93) as a potentially better indicator of central tendency (although mean values were near identical). The corresponding numbers for the control experiment are shown in supplementary Figure S4.1.

A relatively high F_2-score does not automatically imply a skilful forecast, as the control experiment may achieve a similarly high F_2-score. For instance, in forecasting more frequent vegetation impact events ($z(VI)$<-0.5), the control forecast produces a similar median F_2-score of 0.57 (Fig S4-1). Hence, there is little or no skill in the forecasts from this z threshold combination. Overall, the improvements appear to be greatest and most significant for intermediate z-score threshold combinations, increasing by 0.1 to 0.3 F_2-score units when using a threshold of $z(\theta)$<-1.5 to $z(\theta)$<-2.0 combined with $z(VI)$<-1.5 to $z(VI)$<-2.5.

The improvement of *Recall* and *Precision* (Fig. 4.7b and 4.7c) show similar patterns as for the F_2-score across z-score threshold combinations, but the skill was greater in terms of *Recall* (increased by 0.2-0.4) than in *Precision* (increased by 0.1-0.2). These findings suggest that overall skill improvement is primarily due to the reduced number of False Negatives, which dominates recall.

The z-score threshold combinations producing significantly improved skill do not always result in increased lead times. Indeed, decreases in lead time compared to the control forecasts also occurred (Fig. 4.7d). The pattern of improvements across the different forecast metrics is similar for the three soil moisture integration depths. However, θ_{100} showed improvements across a broader range of z threshold combinations.

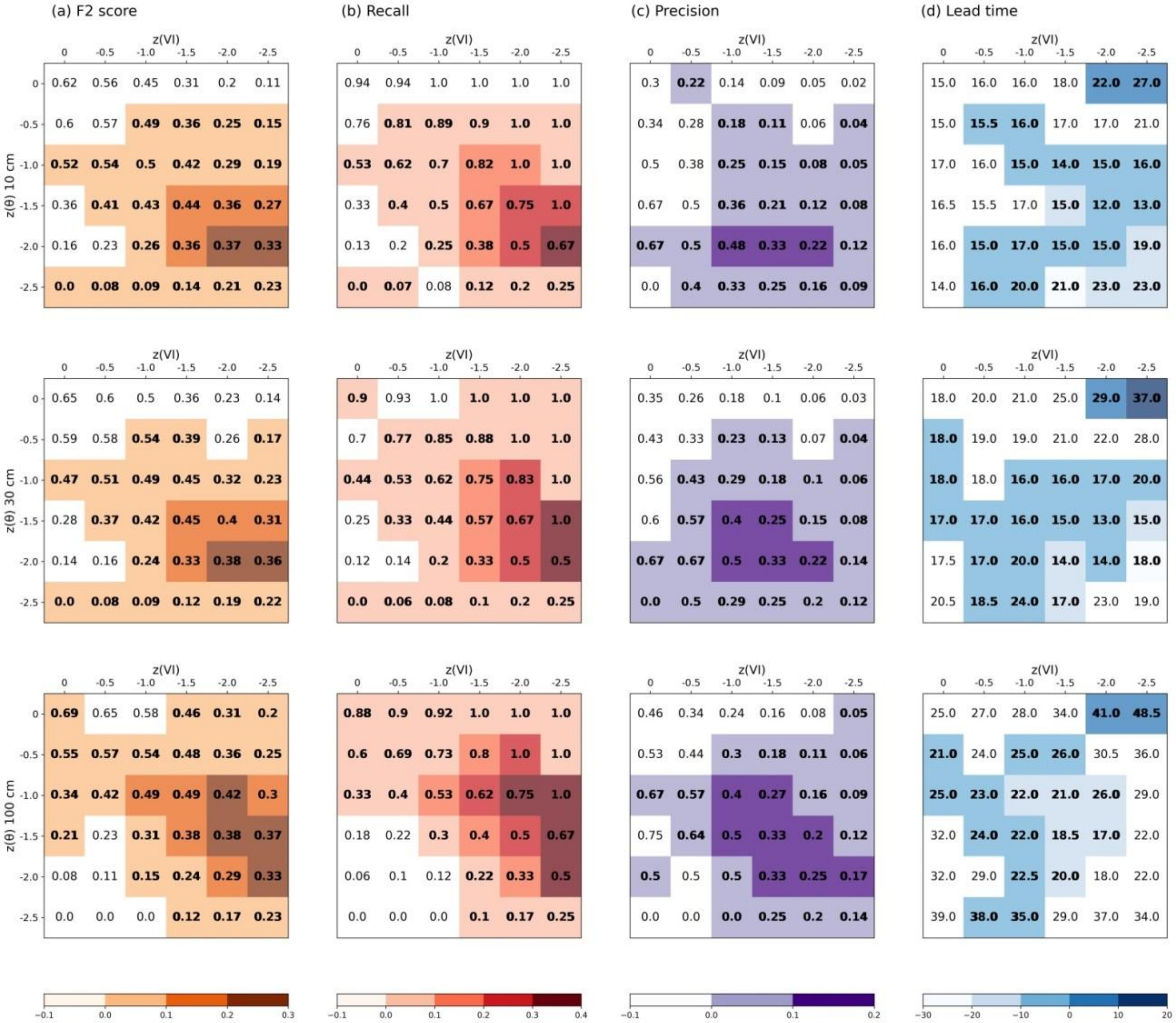


Fig.4.7. The median (a) F_2-score, (b) *Recall*, (c) *Precision*, and (d) *Lead time* of forecasting vegetation threshold event defined by z(*NIRv*) from soil moisture deficiency over three progressive integrated depths (10, 30, and 100 cm) defined by $z(\theta_d)$. Numbers in grid cells represent median values across the 93 in situ sites. The improvement in the metric compared to the control experiment is indicated by the intensity of colour, with bold numbers indicating a significant improvement from the control at p<0.05.

4.3.4. Comparisons between vegetation types and climate regimes

I compared the four predictive ability metrics (F_2-score, *Recall*, *Precision* and *Lead time*) between the identified ecosystems and climate regimes. To limit the scope of the comparison, I focused on VI impacts at z<-1.5. This threshold level was chosen because of the significant skill improvement and the desire to forecast less frequent, more severe impacts. I summarised the improvement of skill score and the achieved forecasting performances and expected *Lead times* in the average of significant z threshold combinations for the three soil layers for different ecosystems and climate regimes (Fig. 4.8).

Savanna vegetation (SAV) showed skill improvement with a relatively narrow distribution of values among sites. The other vegetation types showed lesser F_2-score improvement, and forecast skill improvement was close to zero for woody savanna (WSA) (Fig. 4.8a). Crop-dominant ecosystems (CRO and CVM) showed greater skill improvements and produced the best forecasts at shallow integration depth (θ_{10}). The other vegetation types achieved the best forecast using the greatest integration depth (θ_{100}) (Fig. 4.8b). Considering *Precision* and *Recall* individually helps understand the trade-off between the two in producing the F_2-score. The upper soil layer produced the best F_2-score for crop-dominant ecosystems (CRO and CVM) because of the higher *Precision*, while *Recall* remained similar (Fig. 4.8c and 4.8d). For grasslands (GRA), θ_{100} produced significantly better *Recall* and *Precision* (p<0.05) when compared to shallower integration depths.

Among climate regimes, the greatest skill improvement was found for arid sites, especially when using θ_{100} (Fig. 4.8f). The distribution of actual skill scores appeared similar among dryness categories. Median skill appeared to increase slightly from 0.42 for sub-humid sites to 0.51 for arid sites, but the differences were not statistically significant (p>0.05) (Fig. 4.8g). Similar to vegetation types, *Lead time* increased with integrated soil layer depth, but there was no clear relationship with climate dryness (Fig. 4.8e and 4.8j). Across all climate regimes, the θ_{100} produced higher skill scores than shallower soil with a median *Lead time* of about two to four weeks (12-28 days, Fig. 4.8g and 4.8j). The difference in skill score for different soil depths was greatest for sub-arid sites due to a combination of low *Precision* and high *Recall*. This pattern was weaker for sub-humid sites, where median *Lead time*s were also shorter at around two to three weeks.

In summary, vegetation impact forecasts based on a shallower integration depth produced better forecasts for cropping systems. In contrast, the 0-100 cm soil layer appeared to provide better forecasts than shallower depths for grassland and savanna vegetation types and in sub-arid and arid regimes. Overall poor skill was found for woody savannas.

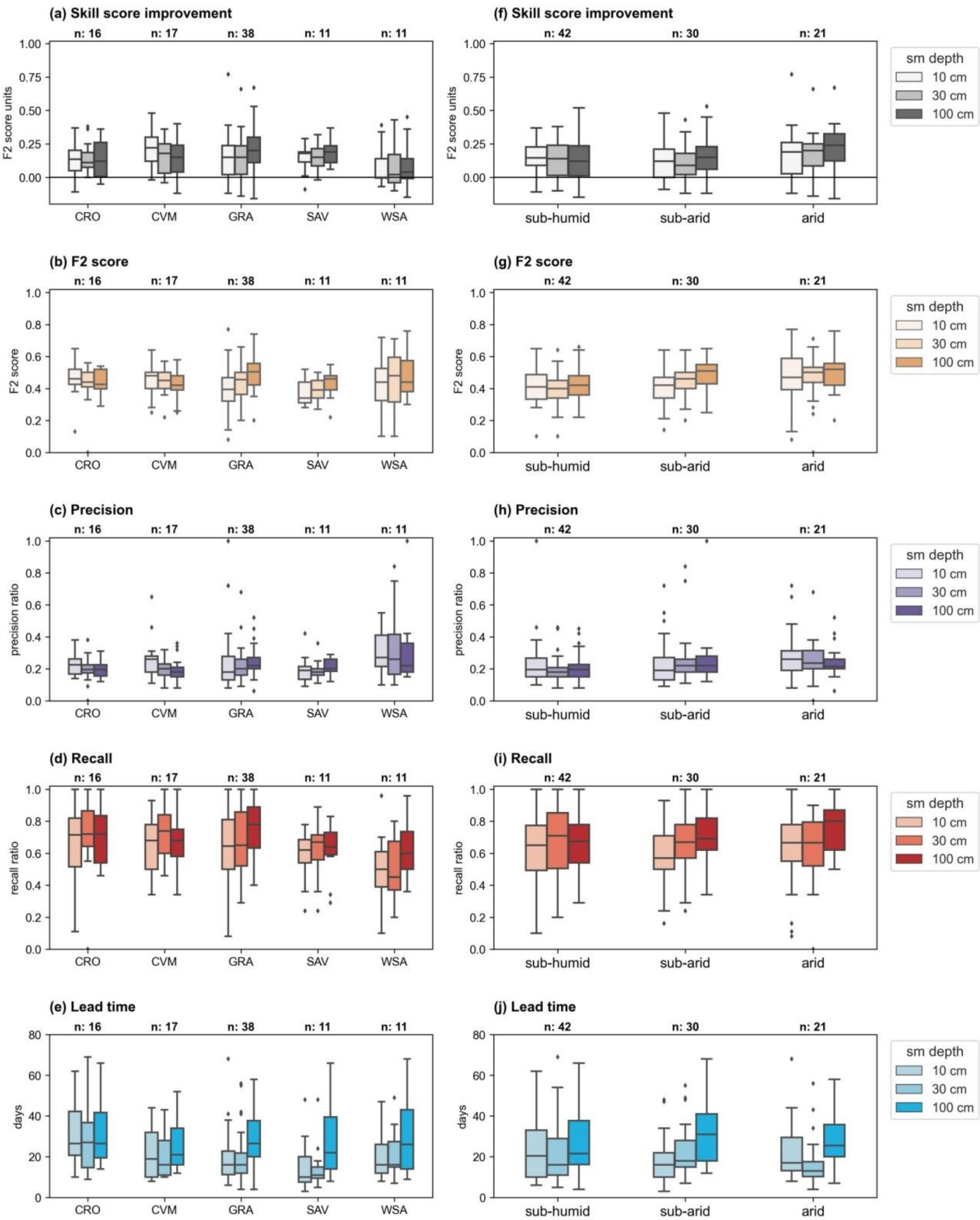


Fig.4.8. Comparison of skill improvement, F_2-score, *Precision*, *Recall* and *Lead time* in forecasting rare and severe VI impact (z(VI)<-1.5) between vegetation types (a-e) and climate regimes (f-j). Grey lines indicate the median values, filled boxes cover the interquartile range, and thin grey lines reach the 5th and 95th percentiles. n denotes the sample size of sites. The outliers are marked as diamonds. CRO: cropland, CVM: crop/natural vegetation mosiacs, GRA: grassland, SAV: savanna, WSA: woody savanna. The dryness classes include sub-humid climate regime, sub-arid climate and arid climate regime.

4.4. Discussion

4.4.1. Limitations and uncertainties

Caveats and uncertainties in my evaluation framework derive from limitations associated with the in-situ soil moisture measurements, satellite-derived VIs, and the forecasting framework applied. I chose to use in situ soil moisture data as I expected that they would offer more accurate estimates of soil moisture over different depths than alternative approaches based on modelling or remote sensing. I was unable to independently ascertain the fidelity of the reported measurements. For the SCAN sites relative errors of ~20% in the recorded data have been reported (Dirmeyer *et al.,* 2016). It is noted that any systematic proportional error or offset would not have affected the calculation of normalised $z(\theta)$ scores, but random errors might. Another limitation of the available in situ soil moisture data is the available record length. By most operational definitions, droughts have a return time of around 25 to 30 years. This definition creates a statistical challenge that affects almost all drought studies because field and satellite records are of similar length at best and typically shorter. The available soil moisture records at root zone depths (e.g., 50 cm and 100 cm) were mostly less than fifteen years. Although the time series did contain one or even more 'true' droughts at some sites, for less rare events, my results should be considered representative of dry events that are infrequent but do not always formally qualify as droughts.

There are also limitations in the satellite vegetation observations used. To determine the occurrence of agricultural drought impacts, I used negative thresholds in standardised z-scores calculated from seasonally-adjusted anomalies in satellite-derived VIs. I avoided VI declines due to phenological and cropping cycles by confining my analysis to within the peak growing season and using seasonally-adjusted anomalies. These adjustments are a deliberate departure from many previous studies. Along with the focus on threshold event forecasting, it creates a much more severe test of forecast skill – one that better represents the true potential merit of an impact forecasting system. A downside of this approach is that the skill metrics and lead time I determined are not easily compared to those reported in previous studies.

Despite the caution in treating VIs, limitations remain. For example, negative VI anomalies during the peak growing season could have occurred for reasons other than soil moisture deficit, such as pests and disease, intensified grazing or harvesting, extreme weather events such as hail or frost, or a change to a crop with markedly different characteristics.

Furthermore, it was mentioned that the VIs were derived from MODIS MCD43A4 temporal composites. In the extreme case, the observation used could represent conditions 8 days earlier or later than the nominal date if observations were unavailable on any of the other 15 days. I expect this problem to have been limited by the prevalence of clear sky conditions during dry periods. The VIs themselves also have limitations in representing true vegetation condition. For example, saturation effects occur in some VIs at high vegetation density, which can reduce their sensitivity to vegetation decline, especially during the peak growing season or if drought only affects the lowest vegetation layers (Xue & Su, 2017).

4.4.2. Merit of the forecasting framework

Past studies of vegetation drought impact forecasting have focused on measuring correlation. This broad focus has limitations when interpreting derived forecast skill and lead time measures in a drought early warning context. Drought impacts are concentrated in one of the tails of the distribution. Forecasts across the full range of conditions can, therefore, easily produce a misleading estimate of the true skill and lead time of drought impact forecasts. By definition, there are, few occurrences of extreme events in the record available to train forecasting models, and this can degrade calculated forecast skill. It also tends to cause an underprediction of severe drought impacts and overprediction of low-intensity drought impacts (Brust *et al.,* 2021; Sutanto *et al.,* 2020). I developed a threshold-based forecasting framework to overcome these issues as best as possible. My framework provides a simple approach to evaluate whether there is any potential in forecasting threshold drought impacts from thresholds in the dry tail of hydro-meteorological predictors. As I have shown, the framework facilitates the selection of optimal predictor and response variables, threshold combinations and other aspects of the forecast model configuration. The framework also enables control or benchmark experiments to evaluate skill and skill improvements – an important feature given the wide variety of available hydro-meteorological and vegetation-related products.

Using the framework developed, I concluded that the standardised anomaly of NIRv marginally outperformed the other satellite-derived VIs. This does not amount to overwhelming evidence that it is a better indicator of vegetation drought impact, although it does support recent studies highlighting the merit of this index. NIRv is thought to respond to the distribution of photosynthetic capacity with depth in canopies, thus more accurately capturing changes in structurally complex landscapes (Badgley *et al.,* 2017). Wang *et al.* (2022) demonstrated that the NIRv might be a better proxy of dryland vegetation productivity dynamics in the western United States. Several alternative VIs could not be considered here.

Future research could consider the Sun-Induced chlorophyll Fluorescence (SIF; Frankenberg *et al.,* 2011), which should more directly relate to photosynthesis, and the Vegetation Optical Depth (VOD; Liu *et al.,* 2011), a measure calculated from satellite microwave radiometry that reflects water status in all above-ground biomass rather than only the leaves visible from above (Liu *et al.,* 2015).

Alternatives are also available for the predictor variable. This study used in situ soil moisture observations at different depths to develop and trial the framework. The application of the framework across larger areas would require spatial soil moisture estimates derived from remote sensing or models such as GLDAS (Rodell *et al.,* 2004) and ERA5-Land (Muñoz-Sabater *et al.,* 2021).Furthermore, using soil moisture forecasts could further extend the lead time for skilful forecasts, depending on how the uncertainty in those soil moisture forecasts propagates into the forecasts of subsequent vegetation impacts.

On theoretical grounds, the preferred combination of z-score thresholds in a drought warning system might be to use identical thresholds for predictor and predictand and reflect rare events, i.e., be strongly negative. These are found along the diagonal near the bottom right of the matrices in Fig. 4.7. The most negative threshold of $z<-2.5$ appeared unsuitable as a $z(\theta)$ trigger, with low F_2-scores resulting from a combination of low *Recall* and low *Precision* (Fig. 4.7a). This is probably due to the small number of true events reaching this threshold at most of the sites. This limitation also reduces the statistical strength of my results for these more severe events. Better combinations do occur along the diagonal for θ_{10} and θ_{30} (Fig. 4.7a), but for θ_{100} a less negative z-score thresholds for soil moisture than for VI appeared to produce better results. The median *Lead time* associated with the corresponding preferred z threshold combinations was 13 to 25 days, i.e. about two to four weeks (Fig. 4.7d).

4.4.3. Forecast skill and the trade-off between missed events and false alarms

The z-threshold combinations producing the greatest improvement in F_2-scores, measured relative to the control simulation, reflect a trade-off between *Recall* and *Precision*, i.e., failing to forecast impacts (missed events) and incorrectly doing so (false alarms). I used the F_β skill score metric to lend greater weight to avoiding missed events. This choice was motivated by the view that missing events would be more problematic than issuing false alarms in drought early warning system (DEWS) (Potter *et al.,* 2021). However, different weightings or skill metrics could be considered and might produce different optimal combinations of predictor, predictand and other aspects of forecast model configuration. Inevitably, in terms of the skill score chosen, reducing the number of missed events (i.e., improving recall) will increase false alarms (i.e., reduce precision). In a DEWS, this would contribute to the 'cry wolf' problem, causing warnings to carry less weight or even be ignored and fail to elicit the necessary response (Kelman & Glantz, 2014). Thus, selecting an appropriate threshold level for the impact that is being predicted is of critical concern (Potter *et al.,* 2021) . My results (Fig. 4.8b and 4.8c) quantify this inevitable trade-off between *Precision* and *Recall*.

An additional dimension of drought forecasting is the possibility of forecasting events of varying severity. I investigated the scope for such a tiered approach by examining decreasing z threshold levels associated with impacts of increasing severity and return time. My results suggest that forecasting relatively more frequent impacts (i.e., $-1.0 \le z < 0$) is associated with high *Recall* and *Precision*, but forecast skill was similar to the control forecast. The relatively high actual and control forecast skill were both due to overprediction. Overprediction is a known issue for interpreting skill scores for an unbalanced sample and one reason why I focused on skill improvements compared to the control experiment (Fig. S4.1). Nonetheless, modest skill improvements still occurred for θ_{100}, suggesting that skilful warning of mild vegetation impacts may be feasible. For impacts of greater severity and lower frequency ($z < -2.0$ and $z < -1.5$), simultaneously avoiding missed alarms and false alarms is more challenging. The significant improvement in skill score compared to the control experiment demonstrated forecasting potential. For the most severe impacts ($z < -2.5$), forecasts were less skilful due to a high number of false alarms, as indicated by low *Precision*. I attribute this to the small number of events in the relatively short site records.

4.4.4. Differences between vegetation types and climate regimes

Different ecosystems and climate regimes sometimes showed differences in optimal θ depths. These differences are assumed to reflect variations in access to soil moisture at different depths (Fig. 4.8). Thus, for cropping systems, θ_{10} appeared optimal for use in early warning. Rooting depths in annual crops can vary dramatically depending on growth stage (Swain *et al.,* 2013). The optimal θ_{10} may have been a compromise for reflecting vegetation response in these complex and dynamically changing ecosystems. The deepest θ_{100} produced the best forecasts for savanna and grassland sites and vegetation in more arid climate regimes. The savanna and grassland sites included in my data mostly comprised native grass species in sub-arid and arid regions (Fig. 4.1). These ecosystems may have adapted to their local environment, giving them the ability to withstand short-term soil moisture anomalies better than rain-fed cropland or pasture species found or introduced in more humid climate regimes to maximise agricultural productivity (Nippert & Holdo, 2015).

Apparently, the more deeply rooted vegetation component was dominant among the sites investigated in this study. The savanna and woody savanna sites showed the least potential for impact forecasting. This finding is consistent with studies by Chen *et al.* (2014) and Nicolai-Shaw *et al.* (2017). They also found that vegetation dynamics in woody savannas have a lower correlation with soil moisture conditions. Tian *et al.* (2019) found that these ecosystems sometimes respond to water availability changes over months rather than weeks, suggesting that soil moisture observations down to 100 cm may be insufficient to forecast drought impacts. Groundwater indices based on water balance models or satellite observations, such as the groundwater resource index (GRI, Mendicino *et al.,* 2008) or GRACE groundwater drought index (GGDI, Thomas *et al.,* 2017) might be more skilful in forecasting changes in savanna and forests ecosystems.

Other studies have also found that soil moisture dynamics in deeper layers are more stable due to the longer memory (Ghannam *et al.,* 2016). Soil moisture thresholds over larger integration depths are less frequently crossed in response to short-term precipitation events or periods that do not saturate the profile. Unfortunately, soil moisture estimates available from remote sensing reflect the moisture content of the top few centimetres of near-surface soil only. Therefore, they are less likely to be skilful in forecasting drought impacts unless used to inform estimates of deeper soil moisture, e.g., through assimilation into land surface models (Tian *et al.,* 2019). The propagation of drought impacts was determined by characteristics such as water-holding capability and seasonal dry-down pathways in different soil profile depths (Sehgal *et al.,* 2020), and vegetation types with varying root-water uptake characteristics. Hence, the analyses of layer-wise drought impact propagation from soil moisture to vegetation could benefit the operation of drought impact forecasting over large spatial scale by generalising the achieved forecast skills by different groups identified based on factors such as soil layers, vegetation types and climate regimes.

Recent attention has been drawn to the relatively fast onset of some drought events, such as the 2012 drought across the central United States, which developed within a few weeks (Otkin *et al.,* 2015; Pendergrass *et al.,* 2020). My results show that vegetation typically responds to soil moisture deficits within two to four weeks. Other studies derived similar results for shallow-rooted vegetation (Asoka & Mishra, 2015; Nicolai-Shaw *et al.,* 2017; Tian *et al.,* 2019). Such a relatively rapid response emphasises the need to develop drought impact forecasting methods that can predict daily or weekly rather than monthly if warnings are to be issued in a timely fashion (Pendergrass *et al.,* 2020).

4.4.5. Implications for early warning and future work

I have found that the lead time between soil moisture triggers and vegetation impacts was generally around two to four weeks. However, these lead times are associated with forecasts based on real-time soil moisture observations. They only consider latency in the availability of past observations. Conversely, they do not consider the additional skill and lead time that may be derived from soil moisture forecasts instead of observations. Seasonal soil moisture outlooks – such as those issued by the USA NOAA Climate Prediction Centre (van den Dool *et al.,* 2003) and the Australian Water Outlook (Vogel *et al.,* 2021) – have been found somewhat skilful for up to three months and therefore may extend lead time further to enable drought preparation several months in advance. Of course, skill inevitably declines with lead time in these forecasts. The expected time lag between soil moisture deficits and vegetation impacts of a few weeks also means that vegetation impact forecasts will be less informative for drought forecasts several months ahead.

Future work could adopt the framework developed here to investigate enhancements and alternative approaches. For example, the in-situ soil moisture observations can be replaced with ensemble soil moisture forecasts (or hindcasts in a research setting) based on a combination of weather forecasting and land surface model(s). This approach can connect the 'rainfall-soil moisture-vegetation' forecasting chain and provide spatially continuous forecasts of vegetation impact.

4.5. Conclusions

I have proposed a framework to forecast the onset of drought vegetation impacts, using in situ soil moisture observations at different depths as the predictor variable and satellite-derived vegetation indices as the predictand. I compared the skill and lead times achieved by forecasts for vegetation impacts of varying severity and across different ecosystems and climate regimes. My main conclusions are as follows:

(1) Standardised soil moisture anomalies could be used to forecast vegetation impacts with significant skill. A challenge in improving forecast skill is the relatively high rate of false alarms when using in situ soil moisture daily observations.

(2) The lead time between soil moisture triggers and subsequent vegetation impacts was typically around two to four weeks. This lead time could be extended using skilful forecasts rather than real-time soil moisture observations.

(3) There is a typical trade-off between recall and precision. Generally, the optimal skill improvements were achieved using similar z thresholds for soil moisture trigger and vegetation impact for moderate severity events. Skill scores were less good for

the most severe events (z-score <-2.5), but this was associated with a small number of events given the limited record length (on average 15 years among the 93 sites).

(4) The soil moisture integration depth that produced the best skill varied between ecosystems, with shallow (0-10 cm) moisture performing better for cropping systems and deeper (0-100 cm) soil moisture being optimal for grassland and savanna as well as for drier environments.

Supplementary Information for Chapter 4

SI Figures

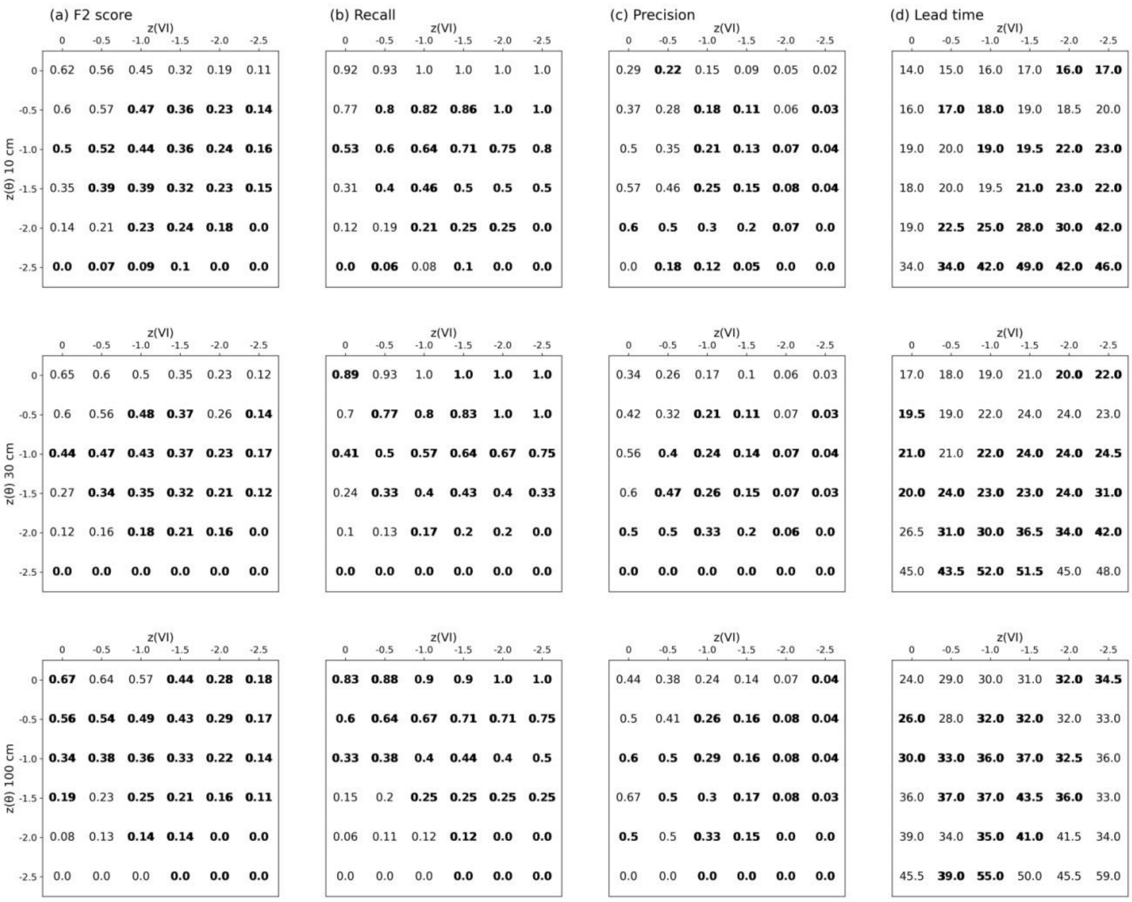


Fig. S4.1. The median (a) F_2-score, (b) *Recall*, (c) *Precision*, and (d) *Lead time* of naïve forecasting achieved from control experiment. The vegetation threshold event defined by z(*NIRv*) and soil moisture deficiency over three progressive integrated depths (10, 30, and 100 cm) defined by $z(\theta_d)$. Numbers in grid represent median values of naïve forecasting across the 93 in situ sites. The bold numbers indicate significant differences from the original experiment at p<0.05.

SI Tables

Table S4.1 Basic information of the selected 93 in situ sites involved for analysis in this study

Site	Latitude	Longitude	IGBP	Temporal coverage	Aridity index	Climate regime	Network
US-IB2	41.84	-88.24	GRA	2005-2018	0.37	sub-humid	Ameriflux
US-Ne3	41.18	-96.44	CRO	2003-2019	0.49	sub-humid	Ameriflux
US-SRG	31.79	-110.83	GRA	2008-2019	0.89	arid	Ameriflux
US-SRM	31.82	-110.87	WSA	2004-2019	0.89	arid	Ameriflux
US-Ton	38.43	-120.97	WSA	2004-2019	0.64	sub-arid	Ameriflux
US-Wkg	31.74	-109.94	GRA	2005-2019	0.91	arid	Ameriflux
ICN-BRW	38.95	-88.96	CRO	2005-2019	0.41	sub-humid	ICN
ICN-DXS	37.44	-88.67	WSA	2005-2019	0.35	sub-humid	ICN
ICN-FAI	38.38	-88.39	CRO	2005-2019	0.38	sub-humid	ICN
ICN-FRE	42.28	-89.67	CVM	2006-2018	0.39	sub-humid	ICN
ICN-FRM	38.52	-89.84	CRO	2005-2019	0.39	sub-humid	ICN
ICN-ICC	40.71	-89.51	CVM	2005-2019	0.39	sub-humid	ICN
ICN-OLN	38.74	-88.10	CVM	2005-2019	0.37	sub-humid	ICN
ICN-ORR	39.81	-90.82	CRO	2005-2019	0.41	sub-humid	ICN
ICN-SIU	37.70	-89.24	CRO	2005-2019	0.35	sub-humid	ICN
OKM-105	34.90	-95.35	CVM	2005-2019	0.43	sub-humid	OKM
OKM-107	36.42	-99.42	GRA	2003-2019	0.76	arid	OKM
OKM-111	34.91	-98.29	CVM	2003-2019	0.66	sub-arid	OKM

Site	Latitude	Longitude	IGBP	Temporal coverage	Aridity index	Climate regime	Network
OKM-113	35.84	-96.00	SAV	2003-2019	0.49	sub-humid	OKM
OKM-120	36.14	-95.45	GRA	2005-2019	0.41	sub-humid	OKM
OKM-13	35.17	-96.63	WSA	2004-2019	0.52	sub-arid	OKM
OKM-15	35.78	-96.35	SAV	2003-2019	0.51	sub-arid	OKM
OKM-19	35.59	-99.27	GRA	2003-2019	0.74	arid	OKM
OKM-20	34.85	-97.00	GRA	2003-2019	0.55	sub-arid	OKM
OKM-23	34.61	-96.33	GRA	2003-2019	0.52	sub-arid	OKM
OKM-32	36.91	-95.89	GRA	2006-2019	0.45	sub-humid	OKM
OKM-33	33.92	-96.32	GRA	2004-2019	0.48	sub-humid	OKM
OKM-34	35.55	-98.04	CRO	2003-2019	0.64	sub-arid	OKM
OKM-36	35.30	-95.66	SAV	2003-2019	0.46	sub-humid	OKM
OKM-38	36.84	-96.43	GRA	2003-2019	0.47	sub-humid	OKM
OKM-39	36.73	-99.14	CVM	2009-2019	0.74	arid	OKM
OKM-44	35.75	-95.64	CVM	2003-2019	0.44	sub-humid	OKM
OKM-56	34.31	-96.00	CVM	2003-2019	0.45	sub-humid	OKM
OKM-59	36.06	-97.21	GRA	2003-2019	0.53	sub-arid	OKM
OKM-6	36.07	-99.90	SAV	2003-2019	0.78	arid	OKM
OKM-61	36.99	-99.01	GRA	2003-2019	0.73	arid	OKM
OKM-65	36.89	-94.84	CRO	2009-2019	0.41	sub-humid	OKM

Site	Latitude	Longitude	IGBP	Temporal coverage	Aridity index	Climate regime	Network
OKM-68	36.90	-96.91	GRA	2003-2019	0.53	sub-arid	OKM
OKM-70	36.74	-95.61	GRA	2003-2019	0.43	sub-humid	OKM
OKM-71	36.03	-96.50	GRA	2003-2019	0.49	sub-humid	OKM
OKM-75	36.36	-96.77	CVM	2003-2019	0.51	sub-arid	OKM
OKM-79	36.36	-97.15	GRA	2003-2019	0.52	sub-arid	OKM
OKM-8	36.80	-100.53	GRA	2003-2019	0.81	arid	OKM
OKM-85	36.42	-96.04	SAV	2006-2019	0.46	sub-humid	OKM
OKM-87	35.54	-97.34	SAV	2003-2019	0.57	sub-arid	OKM
OKM-90	34.88	-96.07	GRA	2003-2019	0.51	sub-arid	OKM
OKM-97	36.78	-95.22	GRA	2003-2019	0.4	sub-humid	OKM
OKM-99	34.98	-97.52	SAV	2003-2019	0.61	sub-arid	OKM
AR-2083	34.25	-92.03	CVM	2004-2006, 2011-2017	0.41	sub-humid	SCAN
AR-2090	35.21	-92.92	CRO	2004-2016	0.39	sub-humid	SCAN
CO-2017	40.87	-104.73	GRA	2005-2019	0.74	arid	SCAN
CO-2197	40.82	-104.71	GRA	2014-2019	0.74	arid	SCAN
IL-2004	40.32	-89.90	CRO	2013-2018	0.4	sub-humid	SCAN
KS-2093	39.79	-99.33	CVM	2007-2019	0.71	arid	SCAN
MN-2050	47.72	-96.27	CRO	2003-2007, 2011-2019	0.45	sub-humid	SCAN

Site	Latitude	Longitude	IGBP	Temporal coverage	Aridity index	Climate regime	Network
MO-2194	37.00	-92.26	SAV	2011-2016	0.41	sub-humid	SCAN
MT-2019	48.48	-109.80	CVM	2003-2019	0.64	sub-arid	SCAN
MT-2117	48.30	-111.92	CRO	2007-2019	0.7	arid	SCAN
MT-2118	48.44	-111.18	CRO	2007-2019	0.73	arid	SCAN
MT-2119	47.06	-109.95	CRO	2007-2019	0.59	sub-arid	SCAN
MT-2121	47.52	-107.13	GRA	2007-2019	0.72	arid	SCAN
ND-2020	46.77	-100.92	GRA	2003-2019	0.57	sub-arid	SCAN
SD-2072	43.74	-96.61	CRO	2005-2019	0.49	sub-humid	SCAN
TX-2105	33.55	-102.37	CRO	2007-2019	0.85	arid	SCAN
UT-2150	41.33	-111.30	GRA	2009-2019	0.46	sub-humid	SCAN
UT-2151	41.34	-111.19	GRA	2010-2019	0.53	sub-arid	SCAN
UT-2160	41.78	-113.82	GRA	2010-2019	0.63	sub-arid	SCAN
UT-2166	38.31	-109.24	GRA	2010-2019	0.62	sub-arid	SCAN
AR_Batesville_8WNW	35.82	-91.78	CVM	2010-2019	0.36	sub-humid	USCRN
AZ_Elgin_5S	31.59	-110.51	GRA	2010-2017	0.82	arid	USCRN
CO_Montrose_11ENE	38.54	-107.69	WSA	2010-2018	0.51	sub-arid	USCRN
CO_Nunn_7NNE	40.81	-104.76	GRA	2011-2019	0.71	arid	USCRN
IA_DesMoines_17E	41.56	-93.29	GRA	2010-2019	0.44	sub-humid	USCRN

Site	Latitude	Longitude	IGBP	Temporal coverage	Aridity index	Climate regime	Network
IN_Bedford_5WNW	38.89	-86.57	CVM	2009-2018	0.32	sub-humid	USCRN
KS_Oakley_19SSW	38.87	-100.96	GRA	2009-2019	0.76	arid	USCRN
MI_Chatham_1SE	46.33	-86.92	WSA	2011-2019	0.43	sub-humid	USCRN
MI_Gaylord_9SSW	44.91	-84.72	WSA	2011-2019	0.41	sub-humid	USCRN
MN_Sandstone_6W	46.11	-92.99	WSA	2011-2019	0.43	sub-humid	USCRN
MO_Salem_10W	37.63	-91.72	SAV	2010-2019	0.36	sub-humid	USCRN
MS_HollySprings_4N	34.82	-89.43	SAV	2009-2019	0.31	sub-humid	USCRN
MT_WolfPoint_29ENE	48.31	-105.10	GRA	2010-2019	0.71	arid	USCRN
NC_Durham_11W	35.97	-79.09	WSA	2010-2018	0.39	sub-humid	USCRN
ND_Northgate_5ESE	48.97	-102.17	CVM	2010-2019	0.54	sub-arid	USCRN
NE_Harrison_20SSE	42.42	-103.74	GRA	2010-2019	0.74	arid	USCRN
NE_Lincoln_11SW	40.70	-96.85	GRA	2010-2019	0.54	sub-arid	USCRN
NE_Whitman_5ENE	42.07	-101.45	GRA	2010-2019	0.67	sub-arid	USCRN
OK_Goodwell_2SE	36.57	-101.61	GRA	2012-2019	0.86	arid	USCRN
OK_Stillwater_2W	36.12	-97.09	CVM	2010-2019	0.53	sub-arid	USCRN

Site	Latitude	Longitude	IGBP	Temporal coverage	Aridity index	Climate regime	Network
OK_Stillwater_5WNW	36.13	-97.11	CVM	2011-2018	0.53	sub-arid	USCRN
SD_Aberdeen_35WNW	45.71	-99.13	GRA	2010-2019	0.54	sub-arid	USCRN
WA_Spokane_17SSW	47.42	-117.53	SAV	2012-2019	0.6	sub-arid	USCRN
WI_Necedah_5WNW	44.06	-90.17	WSA	2010-2019	0.4	sub-humid	USCRN
WY_Sundance_8NNW	44.52	-104.44	WSA	2012-2019	0.56	sub-arid	USCRN

Chapter 5

Early warning of severe vegetation drought impacts based on simulated root-zone soil moisture: an integrated threshold-based framework using copula and event-matching approaches in southeastern Australia

Abstract

Anticipating vegetation drought impacts with sufficient lead time is crucial for proactive drought mitigation. This study evaluates approaches for predicting vegetation deterioration (VD) risks using model-simulated root-zone soil moisture across southeastern Australia, comparing event-matching and copula-based methodologies for multi-categorical severity levels. I develop a threshold-based 'event-matching' framework utilising standardised anomalies of root-zone soil moisture and vegetation indices to identify critical soil moisture trigger thresholds for varying VD severity. The analysis reveals complementary strengths between methodologies: copula-based approaches excel in detecting mild vegetation stress with enhanced false alarm control, while event-matching methods provide more balanced predictions for severe deterioration risks. Agricultural lands demonstrate superior predictive capability compared to natural vegetation, with lead times varying systematically by vegetation type from 0-32 days in agricultural areas to 16-56 days in forested regions. Monthly temporal aggregations significantly enhance prediction capability, improving hit rates from 0.16 to 0.29 for severe impacts while maintaining stable adjacent prediction ratios. Model integration strategies combining copula-based methods for mild stress detection (VD_{1-2}) with event-matching approaches for severe deterioration prediction (VD_{3-5}) demonstrate enhanced capability during the 2002 drought case study. While current lead times provide valuable warning periods, integration with seasonal soil moisture forecasts could extend this prediction horizon. These findings contribute towards developing robust impact-based drought early warning capabilities across extensive geographical regions with diverse vegetation characteristics.

5.1. Introduction

As discussed in Chapters 2 and 3, the slow-onset nature of drought provides a theoretical window for early intervention. However, the development of predictive models for vegetation impacts remains constrained by both statistical and biophysical complexities. The rarity of extreme events and the lack of consistent historical impact records pose major statistical challenges (Vogel *et al.,* 2021). Biophysically, vegetation responses to drought are often non-linear, temporally lagged, and heterogeneous across ecosystems and climate zones, making prediction difficult (Farooq *et al.,* 2012; McDowell *et al.,* 2013; Vicente-Serrano *et al.,* 2013; Zhang *et al.,* 2023). These factors introduce substantial uncertainty into early warning systems that operate across large spatial scales (Berdugo *et al.,* 2020).

Copula-based joint probability modelling has emerged as an effective analytical framework for analysing drought propagation from hydro-meteorological conditions to vegetation response (Tootoonch *et al.,* 2022; Xiang *et al.,* 2023). This approach demonstrates particular strength in capturing tail dependence of low values in variable distributions and provides robust analysis of conditional exceedance behaviour of vegetation responses under soil moisture stress (Zscheischler *et al.,* 2018). However, as demonstrated in Chapter 3, the copula-based methods still face significant limitations in predicting low-frequency yet high-impact vegetation deterioration events.

A logical advance in predicting infrequent vegetation deterioration events involves targeted analysis of discrete occurrences through pattern recognition approaches. Within this broad class, the coincidence analysis has been widely used to quantifies the co-occurrence rate of soil moisture stress and discrete vegetation anomalies within defined temporal windows, is effective for assessing vegetation responses to hydroclimatic extremes across multiple time scales. Applications range from annual (Rammig *et al.,* 2015; Sun *et al.,* 2023a; Yang *et al.,* 2023a; X. Li *et al.,* 2023) to sub-monthly intervals (Tang *et al.,* 2024), including global and mainland China contexts.

In this thesis, the term *event-matching* refers to an empirical, performance-driven framework that identifies threshold combinations with the highest predictive skill for forecasting vegetation deterioration events within predefined early warning windows. Our event matching approach is designed for operational threshold selection rather than statistical analysis of event co-occurrence. Although these studies advance our understanding of biological responses to extreme conditions, current frameworks exhibit limitations in two critical areas:

the practical implementation of multiple ordinal categories and the optimisation of detection thresholds to balance false alarms against missed events in operational warning systems.

My previous research in Chapter 4 developed a novel threshold "event-matching" framework using in-situ soil moisture observations across 93 USA stations (Y. Li *et al.,* 2023). This approach established empirical relationships between soil moisture deficit thresholds and multi-categorical vegetation stress severity levels, highlighting its potential for predicting infrequent but severe events. However, the predictive capability of this framework remains untested for triggering vegetation drought impact warnings across large land masses using model-simulated root zone soil moisture estimates. While these simulated estimates offer expanded geographic and temporal coverage beyond in-situ observations, their inherent uncertainties may affect the balance between false alarms and missed events across regions and severity levels.

This study aims to advance impact-based drought early warning by evaluating approaches for predicting vegetation deterioration (VD) risks across southeastern Australia. I explore the usage of a threshold 'event-matching' framework to identify the soil moisture trigger thresholds for warning multi-categorical VD severity levels. The research has three primary objectives: (1) to test the predictive capability of applying the model-simulated root zone soil moisture over large land mass; (2) to compare 'event-matching' and copula-based approaches to understand their relative strengths in predicting the VD risks across severity levels, particularly for low-frequency but severe events; and (3) through comprehensive performance assessment metrics of prediction skills and lead times, to investigate potential complementary aspects for developing an integrated framework that leverages the strengths of both methodologies across different land use types. This research addresses critical knowledge gaps in understanding how effectively simulated root-zone soil moisture can predict uncommon yet severe vegetation effects, focusing on enhancing prediction capability for events crucial to drought impact management. Due to scarcity of available observed vegetation impacts at a large spatial scale, the persistence prediction of VD risks is used as a reference model for this study.

5.2. Materials

5.2.1. Study area

The study area in southeastern Australia, including Victoria (VIC), southern New South Wales (NSW), and the Australian Capital Territory (ACT), has been introduced in Chapter 3 (see Section 3.3.1). Fig. S5.1 displays Köppen climate classifications, land use types, and spatial skewness patterns of root-zone soil moisture (RZSM) and NDVI anomalies across 2001–2022, while Fig. S5.2 shows land use composition across climate zones, underscoring heterogeneity in land-climate interactions. As noted there, the region spans temperate to semi-arid climates and supports diverse land use types, with dryland cropping and grazing systems dominating (Fig. 5.1a–b). Here, I build on that overview by incorporating additional context from supplementary materials.

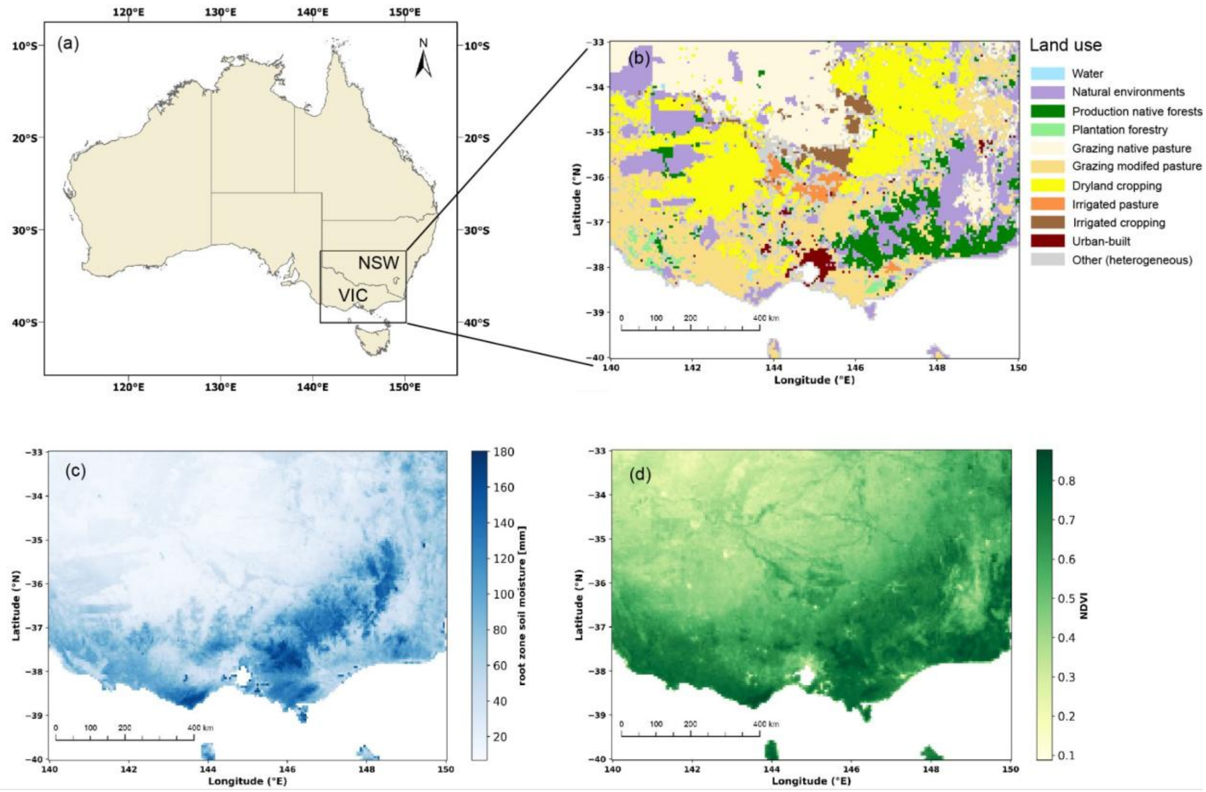


Fig. 5.1. (a) Location map of the study region in Australia (the black square), which encompasses Victoria (Vic), the southern part of New South Wales (NSW), and Australian Capital Territory (ACT). (b) The spatial distribution of major land use types based on reprojected Catchment Scale Land Use of Australia (2018) at 0.05° spatial resolution. The heterogeneous grids are marked in grey colour. The spatial distribution of the mean annual (c) AWRA-L root zone soil moisture and (d) MODIS NDVI over the study area during 2001-2022 peak growing seasons (May to October). The blank grids in the figure are aqua bodies.

5.2.2. Data

The datasets applied in this chapter are consistent with those introduced in Chapter 3 (see Section 3.3.2), including MODIS-derived NDVI, model-simulated root-zone soil moisture (RZSM), and land use classifications based on the Australian Land Use and Management (ALUM) system. Here, I briefly restate the key details relevant to this chapter.

Vegetation condition was measured using the Normalized Difference Vegetation Index (NDVI), derived from the MODIS MCD43A4 product over the period 2001–2022, composited at 8-day intervals and resampled to 0.05° spatial resolution. Only highest-quality reflectance retrievals were retained. The analysis focused on the austral growing season (May-October), coinciding with the primary precipitation period and peak vegetation growth phase across southeastern Australia's wheat-sheep belt.

Root-zone soil moisture was sourced from the Australian Bureau of Meteorology's operational AWRA-L (version 7) model, representing soil moisture content in the 0–1 m profile. The data were averaged to match the NDVI temporal resolution. As described in Chapter 3, the data exhibited substantial right-skewness across many regions. To address this, I applied a square-root transformation to achieve approximate normality; the resulting estimates are hereafter referred to as RZSM.

Land use classifications were drawn from the 2018 Catchment Scale Land Use of Australia dataset and reprojected to 0.05° grid resolution. Following the filtering approach established in Chapter 3, I retained only grid cells where a single land use type accounted for more than 50% of the pixel area, reducing the impact of subgrid heterogeneity on model performance assessments.

5.3. Methods

To evaluate the predictive capability of model-simulated root-zone soil moisture (RZSM) for anticipating multi-categorical vegetation deterioration (VD) severity levels across large spatial scales, this chapter builds upon and integrates two core modelling approaches developed in earlier chapters: (i) the threshold-based event-matching framework (Chapter 4, Sections 4.2.3–4.2.4), and (ii) the bivariate copula method (Chapter 3, Sections 3.4.1–3.4.2).

Both approaches rely on standardised anomalies (z-scores) of RZSM and NDVI, calculated against the 2001–2022 climatology, to characterise the degree of soil moisture deficits and vegetation stress, respectively. For methodological details on z-score derivation, refer to Chapter 3, Section 3.4.1 and Chapter 4, Section 4.2.3. This chapter adopts consistent non-exceedance threshold criteria for VD severity based on standardised VIA thresholds, defining six categories: below-normal ($\leqslant$ 0.0 SD, VD_0), mild ($\leqslant$ -0.5 SD, VD_1), moderate ($\leqslant$ -1.0 SD, VD_2), severe ($\leqslant$ -1.5 SD, VD_3), extreme ($\leqslant$ -2.0 SD, VD_4), and exceptional ($\leqslant$ -2.5 SD, VD_5). Unlike conventional exclusive classification systems with both upper and lower bounds, each VD category is defined by a single upper bound in line with early warning logic of risks that focuses on whether a critical threshold is triggered (WMO, 2016).

This section evaluates the predictive performance of both modelling frameworks using a suite of verification metrics, including Hit Rate, False Alarm Ratio (FAR), and Adjacent Prediction Ratio (APR). Finally, I apply leave-one-out cross-validation to identify optimal integration strategies that combine both models. The integrated framework is assessed using the same performance metrics to ensure robust and operationally relevant predictions across VD severity levels.

5.3.1. Threshold 'event-matching' framework

5.3.1.1. Assessment of potential predictive skill in threshold combination

The threshold event-matching framework, and detailed in Chapter 4, is employed here to identify soil moisture anomaly (SMA) thresholds that precede vegetation index anomaly (VIA) events. This framework applies z-score thresholds to detect threshold-crossing events and quantifies potential predictive performance using F_β-scores across all SMA–VIA threshold combinations. For the underlying methodology, including the event identification logic and key equations, refer to Chapter 4, Section 4.2.3. For practical implementation examples, see Fig. 4.4 in the same chapter.

In this this section, six levels of VIA thresholds are defined ranging from below-normal (VD_0) to exceptional (VD_5), and 30 levels of SMA thresholds spanning from 0 to −3.0 SD at -0.1 SD interval, resulting in 180 SMA-VIA combinations. The predictive skill of each combination is assessed through F_1- and F_2-scores, along with the *Lead time* (Δt) between SMA signals and corresponding vegetation responses. To evaluate whether observed F_β-scores indicate skill beyond random chance, I benchmark them against a bootstrapped random baseline. The baseline is constructed using a temporal randomisation method that preserves spatial coherence and intra-annual structure while disrupting inter-annual dependencies between SMA and VIA. The procedure involves:

1). Randomly shuffling years in the SMA time series to pair each SMA year with a mismatched VIA year;

2). Recalculation of performance metrics (F_1-score, F_2-score, median Δt) based on the shuffled temporal series;

3). Iteration of this process *N* times to generate a distribution of randomised outcomes. The iteration was terminated at $N = 500$, as bootstrapping diagnostics indicated that the mean and variance of the baseline performance metrics (e.g., F_1-score, F_2-score) had stabilised;

4). Computing the mean and median values of these metrics across iterations for each SMA-VIA threshold combination pair.

For each threshold combination at the pixel level, I tested whether the observed (nominal) performance scores significantly exceed those of the random baseline. This was done using Student's *t*-test ($p < 0.05$) to assess statistical significance, with Cohen's *d* used as a complementary criterion to ensure that retained improvements represented a substantial magnitude relative to the stablished random baseline ($d > 0.9$). Threshold combinations satisfying both criteria were therefore considered to demonstrate robust predictive skill beyond random expectations, thereby validating the framework's ability to anticipate vegetation deterioration (VD) events from antecedent soil moisture signals.

5.3.1.2. Trigger threshold determination

The SMA trigger thresholds for each VD severity category were determined by analysing the maximum significant improvements in F_{β}-score across SMA-VIA threshold combinations. To account for the trade-off between missed detections and false alarms, particularly relevant for predicting rare but severe VD events, I applied hierarchical filtering criteria based on both F_1- and F_2-score improvements (i.e., >0.1, >0.2, and >0.3), allowing sensitivity to false alarms and missed events to be explicitly examined during threshold selection.

A hierarchical procedure was applied to derive SMA trigger thresholds across VD severity classes, starting from the most severe VD category and proceeding toward milder VD conditions. While VD severity classes were defined at coarse intervals (0.5 SD), SMA trigger thresholds were searched over a broader and finer range, spanning from −2.9 to 0.0 SD at 0.1 SD increments, to allow precise identification of thresholds associated with improved predictive performance.

To reconcile these differing resolutions and to ensure uniqueness and monotonic progression of SMA thresholds across VD severity levels, a minimum separation constraint of +0.1 SD was introduced. This constraint prevents multiple VD classes from being assigned identical SMA trigger thresholds and avoids non-monotonic ordering that may arise from independent threshold selection across discrete severity categories. For example, under the F_1-score >0.1 filter, only threshold combinations exceeding this benchmark were retained (see Fig. 5.2), and the SMA interval corresponding to the maximum F_1-score improvement was selected for each severity class (Table 5.1). Where no significant improvement was found, no threshold was assigned. This process yielded six sets of SMA trigger thresholds:

EM-f1-0.1, EM-f1-0.2, EM-f1-0.3, EM-f2-0.1, EM-f2-0.2, and EM-f2-0.3—each reflecting different optimisation strategies in predicting VD events.

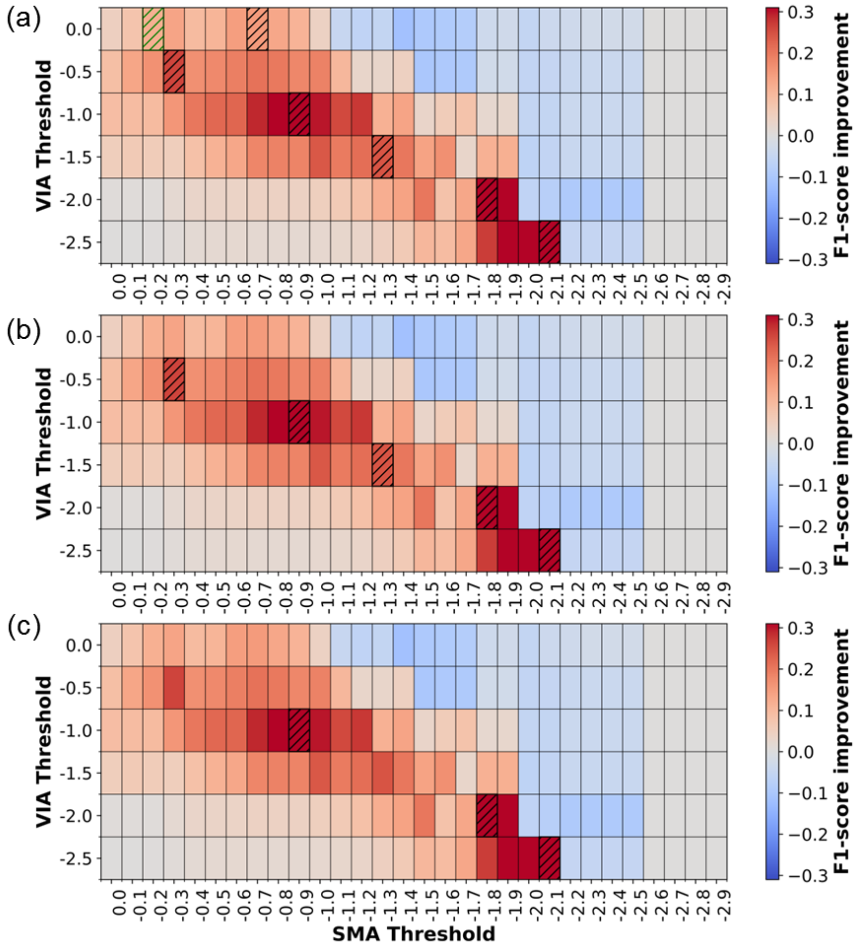


Fig. 5.2. Heatmap visualisations of F_1-score improvements across SMA-VIA threshold combinations in the example pixel [33.75°S, 149.4°E]. Figures (a-c) indicate selection based on different criteria with the improvement of F_1-score >0.1, >0.2, and >0.3, respectively. The x-axis represents SMA thresholds (-3.0 to 0.0 SD at 0.1 increments), whilst the y-axis indicates VD categories (0.0 to -2.5 SD). Colour intensity represents the magnitude of F_1-score improvement relative to the randomised baseline, with red indicating positive improvements and blue showing deterioration in predictive skill. Hatched regions highlight the determined SMA trigger threshold for predicting varying VD severities.

Table 5.1. Example of determined SMA trigger threshold values for varying VD categories in the example pixel [33.75°S, 149.4°E] based on F_1-score improvements.

VD category	VIA threshold	Max F_1-score improvement	SMA trigger with criterion: F_1-score improvement >0.1	SMA trigger with criterion: F_1-score improvement >0.2	SMA trigger with criterion: F_1-score improvement >0.3
VD0	0.0 SD	0.15	-0.2 (-0.7) SD	n.a.	n.a.
VD1	-0.5 SD	0.26	-0.3 SD	-0.3 SD	n.a.
VD2	-1.0 SD	0.34	-0.9 SD	-0.9 SD	-0.9 SD
VD3	-1.5 SD	0.26	-1.3 SD	-1.3 SD	n.a.
VD4	-2.0 SD	0.36	-1.8 SD	-1.8 SD	-1.8 SD
VD5	-2.5 SD	0.63	-2.1 SD	-2.1 SD	-2.1 SD

Note: when SMA trigger with criterion F_1 score improvement >0.1, the SMA trigger for VD_0 is adjusted from -0.7 SD to -0.2 SD to keep the monotonically increasing arrangement.

5.3.2. The bivariate copula method

As introduced in Chapter 3, the copula method estimates conditional probabilities of vegetation deterioration given soil moisture deficits by fitting joint distributions between SMA and VIA. This approach enables threshold estimation based on varying exceedance probability criteria. Readers are referred to Chapter 3, Section 3.4.2 for full theoretical and implementation details.

Following the same setting as in Chapter 3, the framework in this section also employs multiple conditional probability decision criteria $P \in \{\geq 40\%, \geq 50\%, \geq 60\%, \geq 70\%\}$ to establish four estimated sets of critical SMA thresholds, noted as Copula-cp40, Copula-cp50, Copula-cp60, and Copula-cp70.

5.3.3. Evaluation of model predictive performance

5.3.3.1. Vegetation deterioration risk prediction using SMA trigger thresholds

This section establishes a deterministic framework for predicting categorical vegetation determination risk (VD_i) using soil moisture deficit conditions. The methodology employs SMA trigger thresholds$(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$ derived from Section 5.3.1 and 5.3.2 to predict VD risks across multiple severity levels (VD_1 to VD_5). The prediction mechanism employs an indicator function defined as:

$$\mathrm{VD}_{t+L} = \begin{cases} 0, & \text{if}\,\mathrm{SMA}_t > \theta_1 \\ i, & \text{if}\,\mathrm{SMA}_t \leq \theta_i \end{cases}, (i = 1, 2, 3, 4, 5) \tag{5.1}$$

The evaluation incorporates ten different estimate sets of SMA trigger thresholds: six derived through event-matching (i.e., EM-f1-0.1, EM-f1-0.2, EM-f1-0.3, EM-f2-0.1, EM-f2-0.2, and EM-f2-0.3) and four through copula-based analyses (i.e., Copula-cp40, Copula-cp50, Copula-cp60, and Copula-cp70). The framework utilised pixel-specific SMA values at time step t to predict vegetation deterioration severity (VD_i, where $i = 1, 2, 3, 4, 5$) at a future time step $t + L$. Here L represents pixel-specific temporal lag between VIA and SMA quantified through event-matching and copula-base analyses, respectively.

To evaluate the predictive capability of this proposed framework, I applied persistence prediction (PERS) as an unskilled benchmark model, following established practices (Tian *et al.,* 2019; Hao *et al.,* 2017). This benchmark assumes that the vegetation deterioration risk at time step $t + L$ remains unchanged from the current time step t, expressed as $\mathrm{VD}_{t+L} = \mathrm{VD}_t$. As a skill-less reference, PERS serves to contextualise the relative improvements achieved by the proposed approaches.

5.3.3.2. Performance assessment of predictive models

This section implements a suite of evaluation metrics to (i) evaluate the predictive performance of the proposed frameworks and (ii) optimise integrated SMA trigger settings across multiple models to enhance overall VD predictive capability. The evaluation covers all ten SMA threshold sets introduced earlier, including six derived from event-matching framework (i.e., EM-f1-0.1, EM-f1-0.2, EM-f1-0.3, EM-f2-0.1, EM-f2-0.2, and EM-f2-0.3) and four through copula-based analyses (i.e., Copula-cp40, Copula-cp50, Copula-cp60, and Copula-cp70), alongside the persistence prediction (PERS) benchmark. Assessments are conducted at both 8-day and monthly maximum-aggregated scales to account for temporal variability in model performance.

I employ multi-category (6 × 6) contingency tables to analyse the correspondence between predicted and observed vegetation deterioration (VD) categories, which provides detailed insights into how effectively each estimated set of SMA trigger thresholds performs in prediction accuracy. This systematic approach enables comprehensive evaluation of prediction accuracy through examination of the frequency distribution between predicted categories (y_i) and observations (o_i) across different VD categories. Further two types of estimation errors are quantified based on directional deviations in predicted VD risk levels. For each category *i*, I analyse the contingency table elements as follows:

- Hits：correct predictions where predicted category (y_i) equals true category (o_i) (diagonal elements);
- False Alarms：cases where predicted category equals *i* while true category belongs to any other category (sum of elements in *column i,* excluding the diagonal element)
- Misses：cases where true category equals *i* while predicted category falls into any other category (sum of elements in *row i*, excluding the diagonal element)
- Underestimation：cases where predicted category (y_i) is lower than true category (o_i) in category *i.*
- Overestimation：cases where predicted category (y_i) is higher than true category (o_i) in category *i.*

The performance assessment incorporates five verification metrics, calculated at both category-specific and overall levels (Table 5.2):

- Hit Rate (also known as Probability of Detection, PoD): measures the proportion of observed events successfully predicted, particularly suitable for rare event detection.
- False Alarm Ratio (FAR): quantifies the proportion of predicted events that did not occur
- Under-/Overestimation Ratio: indicates the magnitude of prediction bias
- Adjacent Prediction Ratio (APR): serves as a complementary metric alongside over/underestimation ratios to characterise the degree of near-miss.

Table 5.2. Details of category-specific metrics for comparing observed and modelled VD categories.

Metric name	Formula	Range	Ideal value	Notes
Hit Rate (Probability of Detection, PoD)	$\frac{\text{hits}_i}{\text{hits}_i + \text{misses}_i}$	(0,1)	1	Sensitive to hits, but ignores false alarms. Good for rare events (Bennett *et al.,* 2013).
False alarm ratio (FAR)	$\frac{\text{false alarms}_i}{\text{hits}_i + \text{false alarms}_i}$	(0,1)	0	Sensitive to false alarms, but ignores misses (Bennett *et al.,* 2013).
Overestimation Ratio	$\frac{\text{misses}_i\ where\ (\text{y}_i > o_i)}{\text{hits}_i + \text{misses}_i}$	(0, 100%)	0	Useful in quantifying conservative prediction bias.
Underestimation Ratio	$\frac{\text{misses}_i\ where\ (\text{y}_i < o_i)}{\text{hits}_i + \text{misses}_i}$	(0, 100%)	0	Crucial for evaluating the missed warning signals.
Adjacent Prediction Ratio	$\frac{\text{misses}_i\ where\ (\text{y}_i = o_{\text{i}\pm 1})}{\text{misses}_i}$	(0, 100%)	100%	Valuable for indicating reliability of near-miss predictions (Cardoso & Sousa, 2011)

Note: y_i and o_i represent the prediction and observation in category *i*, respectively. *i* denotes the specific VD category being evaluated.

This comprehensive evaluation methodology specifically addresses the challenges inherent in analysing ordinal multi-categorical predictions with imbalanced class distributions, particularly regarding the limited sample sizes in severe deterioration categories (VD_3 to VD_5). Building upon the error evaluation framework of Cardoso and Sousa (2011), I introduce the APR as a metric quantifying the proportion of near-miss predictions. The higher APR values (approaching 100%) indicate that model errors predominantly occur in adjacent categories, suggesting predictions maintain reasonable approximation of true vegetation deterioration intensity despite imperfect prediction. This evaluation enables identification of proposed models that demonstrate significant improvements over baseline persistence prediction, by synthesising accuracy metrics (Hit Rate and FAR) with bias characterisation measures (Underestimation Ratio, Overestimation Ratio and APR) (Bennett *et al.,* 2013; Papagiannaki *et al.,* 2020; Mardian *et al.,* 2023). Detailed example cases demonstrating the calculation of these verification metrics are provided in Appendix Text A1.

5.3.4. Optimisation and cross-validation of integrated SMA trigger threshold strategies

The optimisation framework integrates SMA trigger thresholds derived from copula-based and event-matching approaches to improve prediction across vegetation deterioration (VD) severity levels. Predictive performance over 2001–2022 (*n*=22) is evaluated using a leave-one-year-out cross-validation scheme (Wilks, 2011), in which each year is withheld in turn for validation and model selection is based on performance across all remaining years. Performance is assessed using Hit Rate, FAR, Underestimation Ratio, Overestimation Ratio, and APR.

The selection of optimal integration strategies incorporates category-specific weighted criteria to address varying prediction priorities across the VD spectrum. The weights used reflect both operational priorities and established practice in evaluating rare event prediction models (Murphy, 1993; Wilks, 2011). For lower-severity VD categories ($VD_{1\text{-}2}$), the primary objective is to minimise overestimation, which can lead to unnecessary alarm and resource misallocation. Accordingly, weights are distributed across Hit Rate (30%), FAR (20%), underestimation (10%), overestimation (20%), and APR (20%). The corresponding optimisation objective for lower-severity VD categories is expressed as:

$$Optim_{low} = 0.3 \cdot Hit\,Rate - 0.2 \cdot FAR - 0.1 \cdot Underestim - 0.2 \cdot Overestim + 0.2 \cdot APR \quad (5.2)$$

For higher-severity VD categories ($VD_{3\text{-}5}$), where underestimation poses greater risk for agricultural decision-making, the weights are adjusted to place greater emphasis on capturing severe events. The revised weights prioritise improved detection of rare, high-impact conditions distrusted across Hit Rate (30%), FAR (20%), underestimation (10%), overestimation (20%) and APR (20%). This results in the following optimisation function for higher-severity VD categories:

$$Optim_{high} = 0.3 \cdot Hit\,Rate - 0.2 \cdot FAR - 0.2 \cdot Underestim - 0.1 \cdot Overestim + 0.2 \cdot APR \quad (5.3)$$

These weights were chosen to reflect a balance between sensitivity and reliability and are consistent with prior frameworks for both deterministic and probabilistic model evaluation (Wilks, 2011), while also aligning with operational goals of impact-based early warning systems.

5.4. Results

5.4.1. Distribution of vegetation anomalies by year

Fig. 5.3 shows the distributions of VIA for the 18,086 study grid cells featuring cropland and pasture during the peak season for 2001-2022. The largest number of grid cells with values below zero occurred for the seasons in 2002, 2006, and 2018. My study period includes the “Millennium Drought ", which was from 2001 to 2009 over southeastern Australia (van Dijk *et al.,* 2013), showing years with worse and less bad drought conditions in terms of negative z-score values. The long low tail in 2020 is caused by forest burning during the unprecedented Black Summer wildfire in southeastern Australia (Levin *et al.,* 2021).

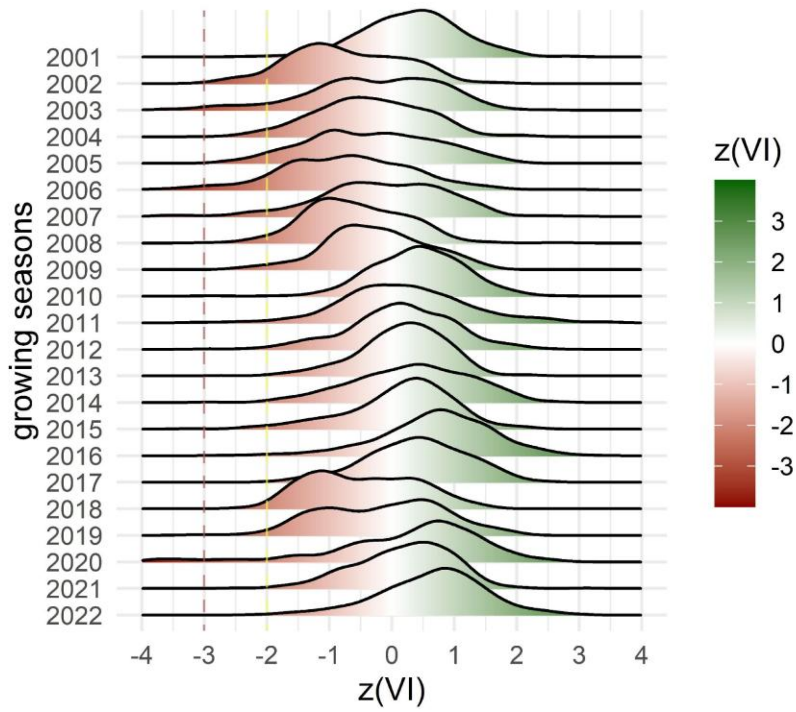


Fig. 5.3. Histogram for the anomalies of vegetation condition in agricultural land for the 22 growing seasons between 2001 and 2022 and for the 18,086 pixels selected.

5.4.2. Characteristics of vegetation deterioration (VD) onset events

This section characterises the frequency and spatial distribution of vegetation deterioration (VD) onset events across southeastern Australia. Here, VD onset events are defined as the first annual occurrence at which the VIA reaches a given severity threshold (VD_0–VD_5). Subsequent threshold crossings within the same year are not counted, thereby avoiding repeated event identification due to short-term oscillations around threshold values.

The frequency of VD onset events shows no pronounced spatial clustering at the grid-cell level. However, the frequency of VD onset events generally increases from inland toward coastal regions, broadly aligning with the spatial distribution of land-use types (Fig. 5.4a-d). For VIA severity thresholds ranging from -1.0 to -2.5 SD, corresponding to moderate to exceptional VD severity, event frequencies vary from approximately 1 to 0.12 events per year across land-use categories, equivalent to return intervals of roughly 1 to 8 years (Fig. 5.4f-j). In subsequent analyses, particular attention is given to events reaching VIA ≤ −1.5 SD, including more extreme thresholds, corresponding to moderate-to-severe vegetation deterioration. VD onset events at these severity levels occur less than once per year across all land-use types.

Plantation forestry exhibits the highest frequency of vegetation deterioration (VD) onset events, likely reflecting the sensitivity of remotely sensed vegetation indices to forest management activities such as harvesting-related canopy disturbances. This is followed by grazing on modified pasture and irrigated pasture. Dryland cropping and irrigated cropping show frequencies comparable to those observed in natural environments and production native forests, while grazing on native pasture exhibits the lowest frequency of VD onset events.

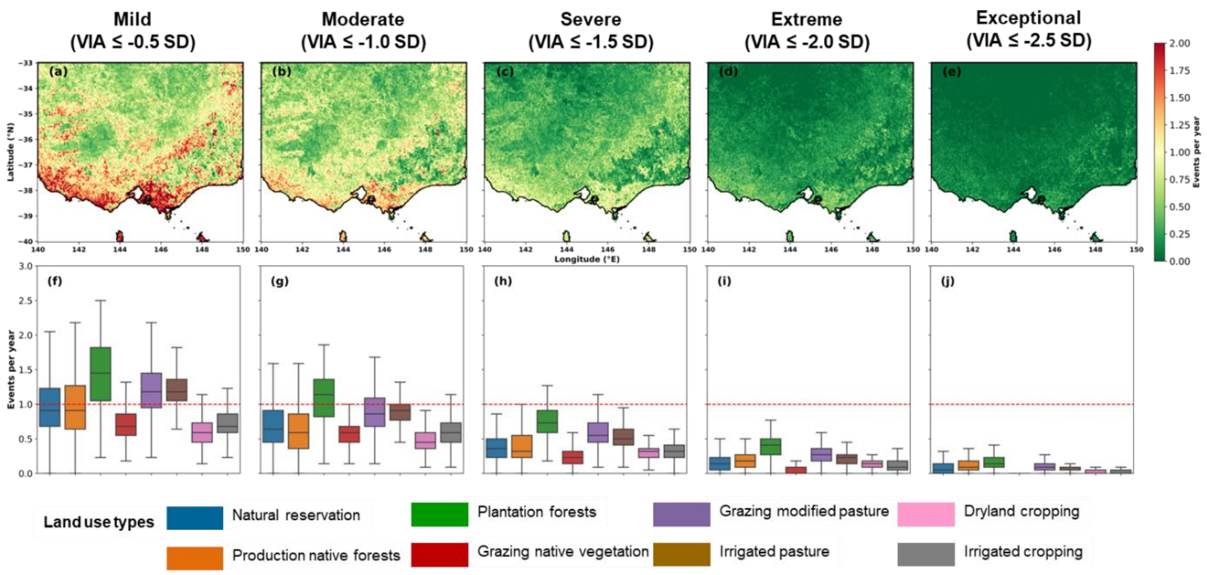


Fig. 5.4. Panels (a-e): Spatial distribution of the mean annual frequency (events yr^{-1}) of identified vegetation index anomaly (VIA) threshold-crossing events during the growing season (May–October) from 2001 to 2022, across five vegetation deterioration (VD) severity levels: (a) mild (VIA ≤ –0.5 SD), (b) moderate (≤ –1.0 SD), (c) severe (≤ –1.5 SD), (d) extreme (≤ –2.0 SD), and (e) exceptional (≤ –2.5 SD). Panels (f-j): Boxplots showing the distribution of mean annual frequency of corresponding VIA threshold events across different land use types. The red dash lines indicate the frequency threshold of one event per year.

5.4.3. Variation in F_β-score and lead time among threshold combinations

Different combinations of [SMA, VIA] thresholds yield varying degrees of significant improvement in F_β-score (Fig. 5.5) and *Lead time* (Fig. S5.3). Overall, the strongest significant improvements in both F_1- and F_2-scores are concentrated near the diagonal line where SMA and VIA thresholds are similar (Fig. 5.5, left column), suggesting that pairing similar thresholds may enhance the reliability of vegetation drought impact prediction, especially as impact severity increases. I find that the most substantial skill improvements, ranging from 0.1 to 0.3 units on the F_2-score, occur when using thresholds of SMA ≤ -1.0 to ≤ -2.0 SD to predict moderate to extreme VD events (VD_2 to VD_4). Comparing Fβ-score improvements across land use types (Fig. 5.5, middle and right columns) reveals stronger predictive performance in agricultural land uses for severe to extreme events (VD_3 to VD_4). In contrast, non-agricultural areas show lower predictive skill.

Although some threshold combinations yield significantly improved F_β-scores, they do not consistently correspond to significantly extended *Lead times* compared to the bootstrapped baseline (Fig. S5.3 bottom row). Here, a positive Lead time is defined as SMA preceding VIA. Longer *Lead times* tend to be associated with the prediction of more severe vegetation drought impacts (Fig. S5.3 top row). For instance, ranging from approximately 24 to 48 days for extreme events (VIA ≤ -2.0 SD), compared to 16 to 32 days for moderate conditions (VIA ≤ -1.0 SD).

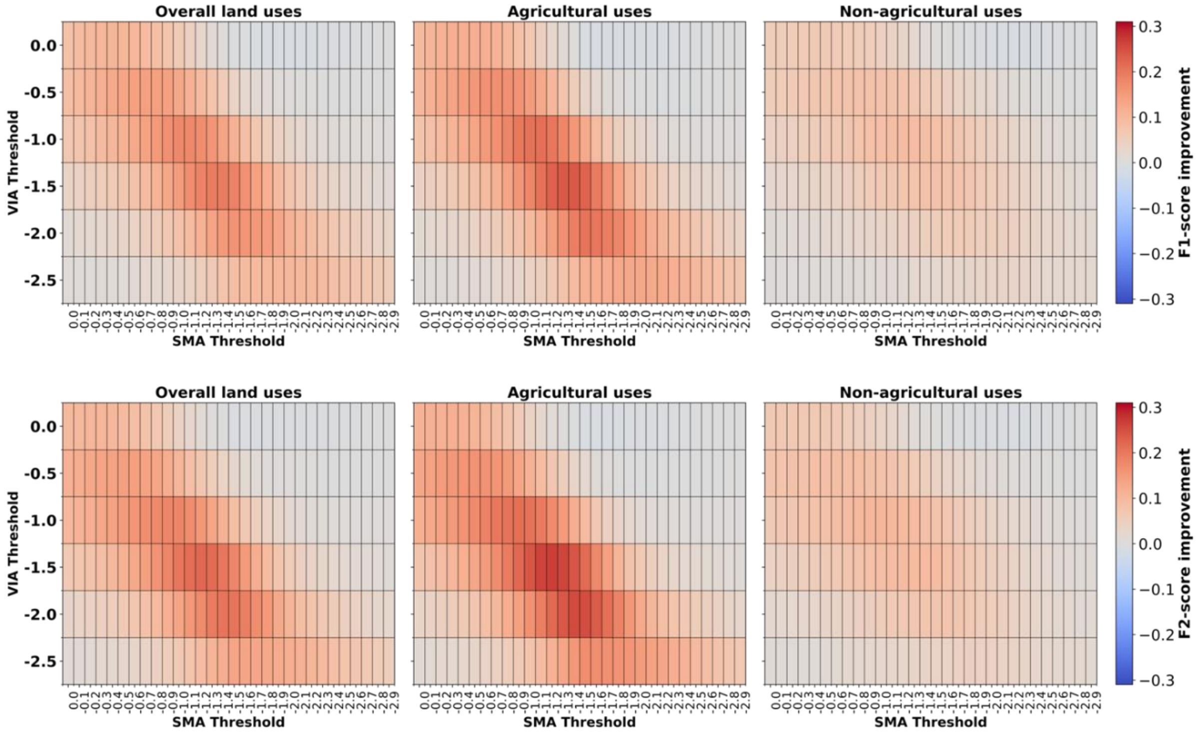


Fig. 5.5. Improvement in F_1-score (top row) and F_2-score (bottom row) for each combination of vegetation index anomaly (VIA) and soil moisture anomaly (SMA) thresholds, relative to a bootstrapped random baseline. Scores shown here reflect threshold-based predictive performance only and do not incorporate lead-time information. Results are summarised by land use types: overall, agricultural and non-agricultural. For baseline generation methodology, see Section 5.3.1.3.

5.4.4. SMA trigger thresholds determined through the event-matching framework

This section focuses on SMA trigger thresholds derived from the event-matching framework. Copula-based thresholds and their spatial characteristics are presented in Chapter 3 and are subsequently incorporated into the integrated threshold analysis in Section 5.5.

Six 'event-matching' (EM-based) variants were developed, based on two evaluation metrics (F_1-score and F_2-score) with three significance criteria settings (improvements ≥ 0.1, 0.2 and 0.3). The full sets of estimated SMA trigger thresholds are presented in the supplementary materials (Fig. S5.4 and S5.5). The event-matching framework revealed clear spatial patterns in SMA trigger thresholds across varying F_β-score improvements and associated VD severity levels. Both F_1- and F_2-based thresholds exhibited a pronounced coastal-inland gradient in trigger values, with eastern regions characterised by pastures requiring more severe soil moisture stress (depicted in darker red) to trigger unfavourable vegetation conditions. Regions associated with dryland cropping tend to exhibit less negative SMA trigger thresholds, corresponding to lighter colouration and indicating sensitivity to relatively modest soil moisture deficits. As VD conditions intensified from -0.5 SD to -2.5 SD, pixel colouration intensified correspondingly, reflecting the increased soil moisture deficit severity required for triggering more severe VD events. This relationship was particularly pronounced under severe VD conditions (-2.0 to -2.5 SD) in eastern forested regions.

Statistical results indicated that SMA trigger thresholds derived from the F_1-score were consistently lower (i.e., more negative) than those based on the F_2-score (Table 5.3). Across both scoring methods, increasing the improvement criterion from 0.1 to 0.3 produced a systematic shift toward more negative trigger values. This stricter threshold also reduced the spatial coverage of identified SMA triggers, with the most pronounced reductions observed in southern and inland grazing regions.

Table 5.3. Summary of event-matching estimated sets of SMA trigger threshold value for varying vegetation deterioration risks.

VD risk category	EM-f1-0.1	EM-f1-0.2	EM-f1-0.3	EM-f2-0.1	EM-f2-0.2	EM-f2-0.3
VD_1	-0.8 [-1.0, -0.5]	-0.8 [-1.0, -0.6]	-0.8 [-0.9, -0.6]	-0.6 [-0.8, -0.3]	-0.6 [-0.8, -0.4]	-0.6 [-0.8, -0.4]
VD_2	-1.1 [-1.3, -0.9]	-1.2 [-1.3, -0.9]	-1.2 [-1.3, -1.0]	-0.9 [-1.1, -0.7]	-0.9 [-1.1, -0.7]	-1.0 [-1.1, -0.8]
VD_3	-1.4 [-1.6, -1.2]	-1.5 [-1.6, -1.3]	-1.5 [-1.6, -1.3]	-1.2 [-1.5, -1.0]	-1.3 [-1.5, -1.1]	-1.3 [-1.5, -1.1]
VD_4	-1.7 [-1.9, -1.4]	-1.8 [-1.9, -1.5]	-1.8 [-2.0, -1.5]	-1.5 [-1.8, -1.3]	-1.6 [-1.8, -1.4]	-1.6 [-1.8, -1.4]
VD_5	-2.0 [-2.3, -1.6]	-2.0 [-2.3, -1.7]	-2.1 [-2.3, -1.7]	-1.8 [-2.1, -1.5]	-1.9 [-2.2, -1.6]	-1.9 [-2.2, -1.6]

Notes. The values of SMA trigger are presented as mean value (SD) with interquartile range [25^{th} percentile, 75^{th} percentile]

The event-matching framework produced consistent median *Lead times* across both F_β-score series and improvement criteria. Identified *Lead times* progressively increased with VD severity. These forecast horizons ranged from 0 days for mild deterioration (VD_1) to 32 days for exceptional conditions (VD_5), with intermediate lead times of 8 days, 16 days, and 24 days observed for VD_2, VD_3, and VD_4 respectively (Table S5.1). Lead times consistently increase with VD severity, indicating greater potential early warning horizons for more severe vegetation deterioration events.

5.4.5. Capability in predicting vegetation deterioration (VD) severity categories

5.4.5.1 Performance across VD severity categories

Figure 5.6 summarises the predictive performance of copula-based and event-matching approaches across VD severity categories at both 8-day and monthly temporal resolutions. Figure 5.6a shows Hit Rates, while Fig. 5.6b presents the corresponding False Alarm Rates (FAR), enabling direct assessment of recall–specificity trade-offs across modelling approaches. Across all approaches, no-deterioration conditions (VD_0) are consistently well predicted, with high Hit Rates (0.82–0.94) and relatively low FARs. Predictive skill declines sharply for mild deterioration categories (VD_1–VD_2) across all approaches. In these categories, differences between copula-based and event-matching approaches remain modest, while the persistence benchmark (PERS) achieves higher Hit Rates but at the cost of substantially elevated FARs.

Clear divergence emerges for severe deterioration categories (VD_3–VD_5). As shown in Fig. 5.6a, EM-based approaches achieve markedly higher Hit Rates than copula-based thresholds at both temporal resolutions, with further improvement under monthly aggregation. However, Fig. 5.6b indicates that this improved recall is accompanied by increased FARs, particularly for EM-based strategies. In contrast, copula-based approaches remain highly conservative, exhibiting very low FARs but correspondingly poor recall for severe VD events. Temporal aggregation from 8-day to monthly resolution modestly improves hit rates and reduces FAR, particularly for EM-based models. Together, these results highlight a fundamental trade-off between conservative detection and effective identification of severe vegetation deterioration.

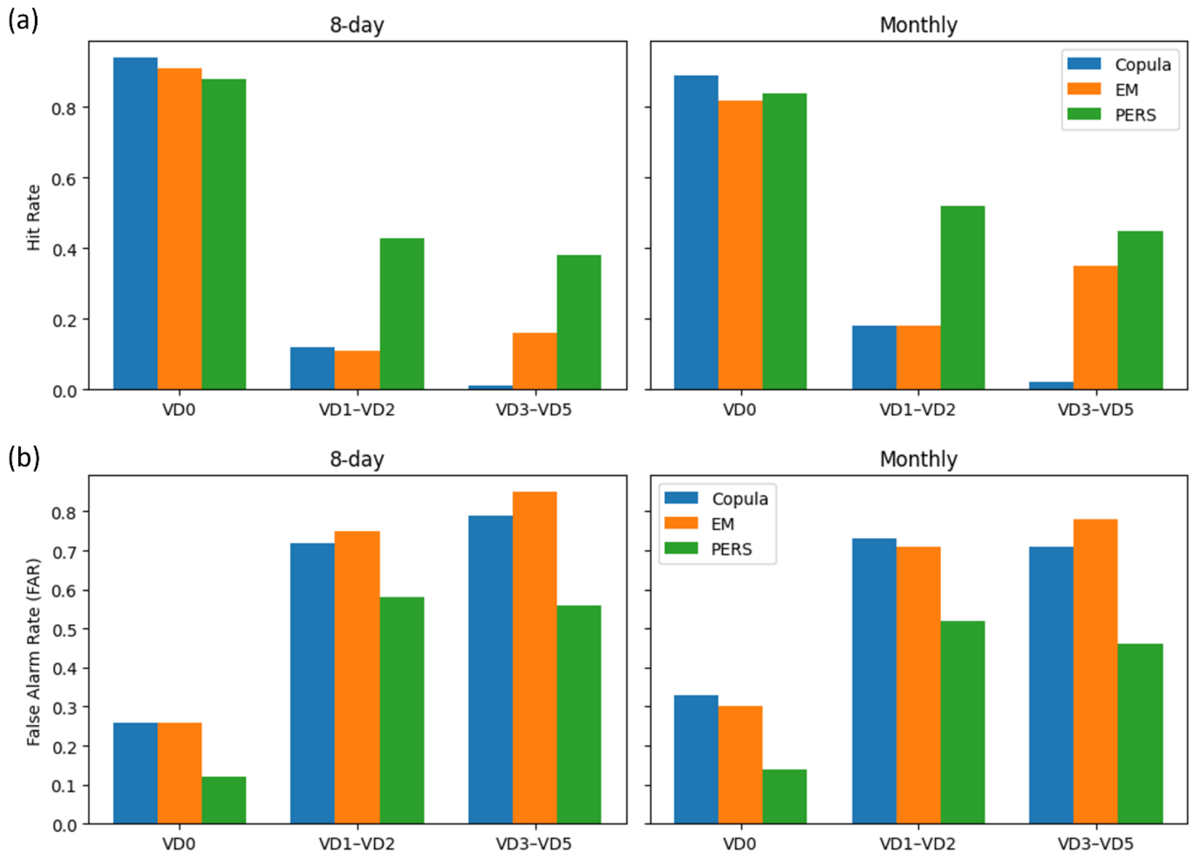


Fig. 5.6. Predictive performance across vegetation deterioration (VD) severity groups for copula-based, event-matching (EM)–based, and persistence (PERS) approaches at (a) 8-day and (b) monthly temporal resolutions. Panel (a) shows Hit Rates and panel (b) shows False Alarm Rates (FAR), aggregated for VD_0 (no deterioration), VD_1–VD_2 (mild deterioration), and VD_3–VD_5 (severe deterioration) groups. For copula-based and EM-based approaches, values represent approach-level averages across multiple SMA trigger sets, based on overall performance statistics reported in Tables S5.2 and S5.3.

5.4.5.2 Optimal performance requirements

To support the design of an effective early warning system (EWS) for drought-induced vegetation deterioration, this study defines optimal performance requirements that account for the ordinal structure of impact severity and their relative importance. The primary objective centres on achieving high detection accuracy for severe events whilst maintaining minimal false alarms for low-risk conditions. Based on these considerations, I establish the following key requirements for optimal predictive performance (with reference to the "perfect score" benchmarks defined in Table 5.2):

1. High hit rate for severe deterioration events (VD_3-VD_5) — aiming toward a perfect score of 1.0

2. Minimal underestimation in VD_4-VD_5 categories and adjacent prediction ratio (APR) ideally approach 100% for any misclassified among high-risk events.

3. Low false alarm rate (FAR) in VD_1-VD_2 predictions, accepting moderate hit rate reduction.

4. Minimal overestimation in the VD_0-VD_2 range and APR ideally at or near 100% for frequent, low-risk scenarios.

5. High total APR across all severity levels — with a reference target of 100%.

Table 5.4 presents the assessment of optimal performance requirements across model configurations at the 8-day temporal scale. Overall, none of the candidate models satisfies all five requirements simultaneously, highlighting the difficulty of achieving balanced predictive performance across VD severity levels. Among the EM-based configurations, EM-f2-0.1 satisfies two of the five requirements, demonstrating relatively strong performance in severe event prediction (*hit rates:* VD_3=0.16, VD_4=0.17, VD_5=0.25) and reduced underestimation for high-risk scenarios (*APR:* VD_4=72%, VD_5=75%). The copula-based models (Copula-cp40 to cp70) share the same fitted copula types but differ by conditional probability decision criteria. Copula-cp50 exhibits comparatively lower false alarm rates in low-to-moderate categories (*FAR:* VD_1=0.71, VD_2=0.7), while Copula-cp70 shows minimal overestimation in low-risk scenarios (*overestimation*: VD_0=1%, VD_{1-2}=0%). In terms of overall classification consistency, Copula-cp40 achieves the highest total adjacent prediction ratio (APR = 43%) among the copula-based configurations.

The temporal aggregation from 8-day scale (Table S5.2) to monthly scale (Table S5.3) demonstrates notable variation in predictive performance. For frequent but slight deterioration (VD_{1-2}), monthly aggregated predictions show enhanced capability with hit rates improving from 0.33 to 0.46 for VD_1 in Copula-cp40, while maintaining similar false alarm rates. The event-matching (EM) models exhibit notably robust performance at the monthly scale for severe deterioration prediction (VD_{3-5}), with EM-f2-0.1 achieving hit rates of 0.29, 0.34, and 0.58 for VD_3, VD_4, and VD_5 respectively, compared to 0.16, 0.17, and 0.25 at the 8-day scale. This enhancement is particularly evident in extreme vegetation deterioration predictions (VD_{4-5}), where the monthly scale reduces underestimation errors. However, this improvement comes with a trade-off in overestimation, which increases from 15% to 24% in the overall performance for EM-f2-0.1. These findings suggest that monthly maximum aggregation may be advantageous for warning of sustained severe vegetation stress, despite the trade-off in temporal resolution.

Table 5.4. Performance requirements for vegetation deterioration (VD) predictability and best-performing SMA trigger threshold sets at 8-day temporal scale.

Best-performance Requirements	Best set of SMA trigger thresholds	Best performance metrics	Perfect score
1: High hit rate for severe deterioration (VD_3-VD_5)	EM-f2-0.1	VD_3: 0.16, VD_4: 0.17, VD_5: 0.25	1.0
2: Minimal underestimation in VD_4-VD_5 with high APR	EM-f2-0.1	VD_4: 72% (26%), VD_5: 75% (7%)	0% (100%)
3: Low FAR in VD_1-VD_2 predictions	Copula-cp50	VD_1: 0.71 VD_2: 0.7	0.0
4: Minimal overestimation in VD_0-VD_2 with high APR	Copula-cp70	VD_0: 1% (93%) VD_1: 0% (100%) VD_2: 0% (10%)	0% (100%)
5: High total APR	Copula-cp40	Overall: 43%	100%

Note: Values in parentheses denote the complementary misclassification metric associated with the primary performance indicator (e.g. underestimation or adjacent prediction ratio, APR).

5.4.6. Enhanced VD prediction through model integration: Case study

All results presented up to this point are derived from the leave-one-year-out cross-validation (LOOCV) framework described in Section 5.3.4. Building on these validated results, this subsection examines the practical performance of the integrated modelling strategy using representative case studies.

The model integration strategy outlined in Section 5.3.4 was implemented to optimise VD prediction across agricultural systems under the LOOCV framework. The optimised configuration utilised Copula-cp40 SMA trigger thresholds for lower-severity VD categories ($VD_{1\text{-}2}$) and EM-f2-0.1 trigger threshold set for higher-severity VD categories ($VD_{3\text{-}5}$), enabling improved representation of both mild and severe vegetation deterioration at 8-day and monthly temporal scales. Fig. 5.7 shows the spatial distribution of key agricultural land-use types and identifies the three case study regions used to evaluate the integrated predictive framework. For each region, long-term seasonal VD patterns over 2001-2022 are presented in Fig. S5.6.

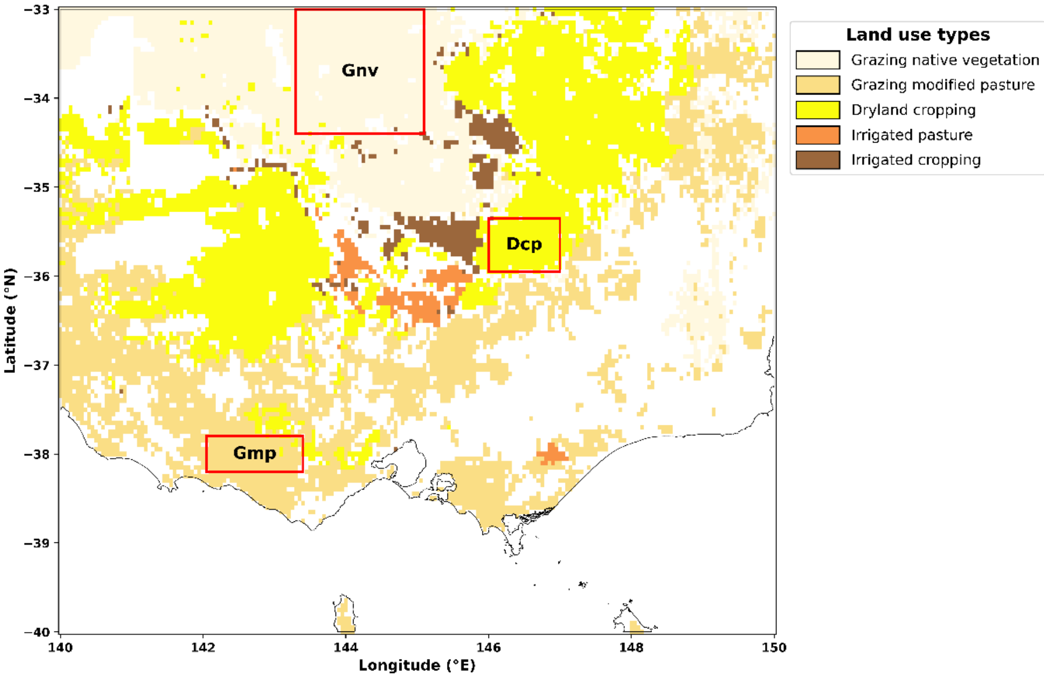


Fig. 5.7. Spatial distribution of major agricultural land use types in southeastern Australia, with case study regions outlined in red: dryland cropping (Dcp), grazing modified pasture (Gmp), and grazing native vegetation (Gnv).

Southeastern Australia experienced one of the most severe droughts on record during 2002-2003 (van Djik, 2013). I examine the year 2002 as the primary case study (Fig. 5.8), focusing on dryland cropping (Dcp) regions where this rapidly developing drought caused substantial agricultural losses (ABARES, 2020). Additional case studies covering grazing modified pasture (Gmp, 2008), grazing native vegetation (Gnv, 2018), and non-drought years are presented in the supplementary appendix (Fig. S5.7 to Fig. S5.9). These selected subregions comprise approximately 95% single land use classification pixels.

Fig 5.8 compares monthly-scale VD predictions from the event-matching model (EM-f2-0.1), the copula-based model (Copula-cp40) and the integrated framework against benchmark predictions and observed VD conditions. For the 2002 growing season, the integrated model demonstrates improved predictive capability in dryland cropping systems, despite a limited lead time of 0–1 month. In particular, the integrated framework shows closer agreement with observations (Fig. 5.8b), capturing both the temporal progression of vegetation stress from May to September (left panel) and the spatial distribution of VD severity categories (right panel). It successfully predicts that moderate to severe deterioration (VD_3–VD_5) affects approximately 60% of the study area, while limiting overprediction of mild deterioration (VD_1–VD_2).

In contrast, predictive performance for grazing systems is substantially weaker. Both grazing on modified pasture (Gmp) and grazing on native vegetation (Gnv) exhibit lower hit rates for severe VD categories and less consistent temporal evolution compared with cropping systems, under the same integrated framework. These results are documented in the supplementary case studies (Figs. S5.7–S5.9) and indicate that the benefits of model integration are land-use dependent. Overall, the cropping case study illustrates the potential of the integrated approach to enhance drought-induced vegetation deterioration prediction using root-zone soil moisture at large spatial scales, while highlighting its more limited performance in grazing systems.

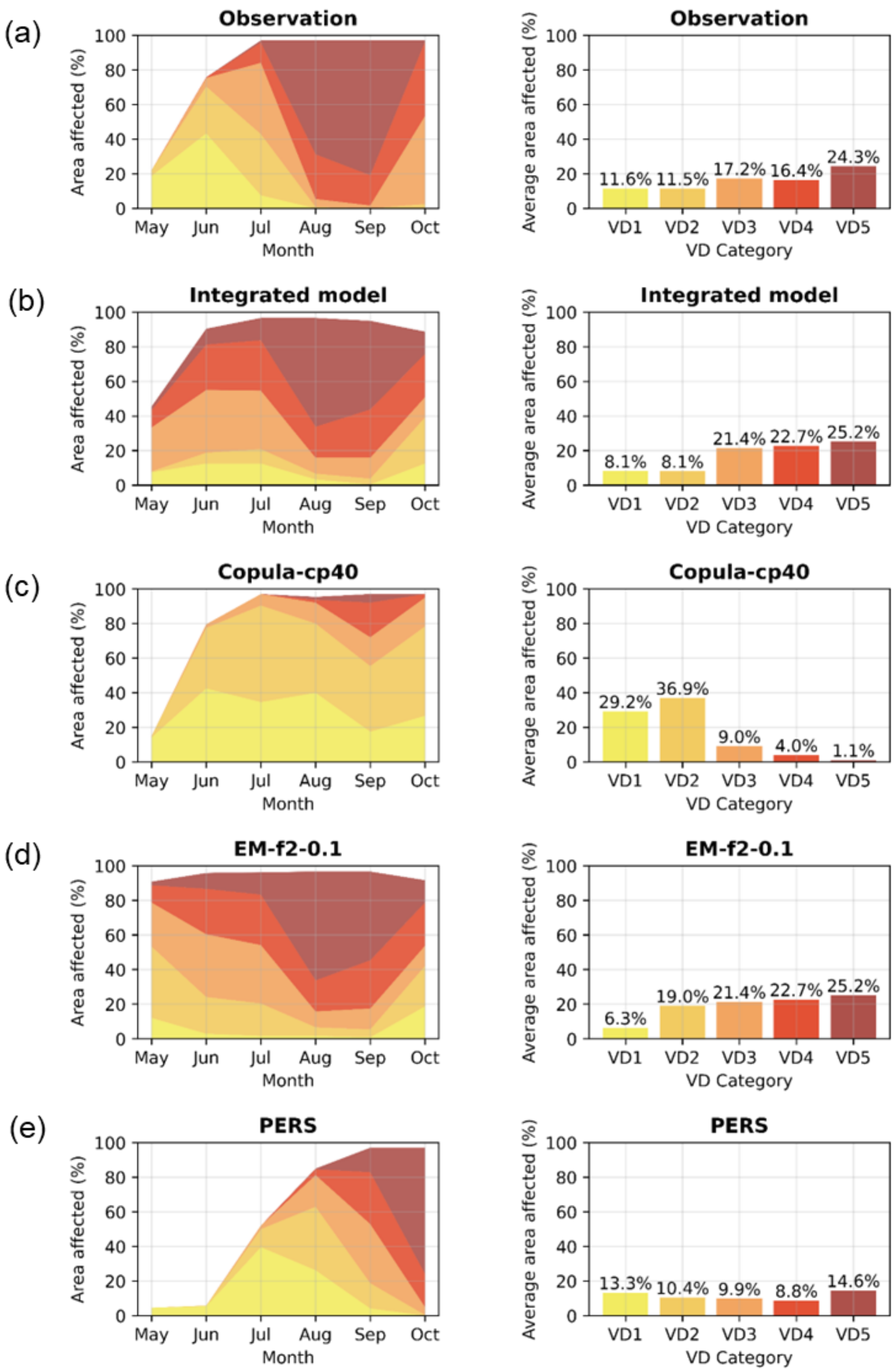


Fig.5.8. Comparison of vegetation deterioration (VD) predictions from different models for the dryland cropping (Dcp) region during the 2002 drought. Left panels show the monthly progression of the area (%) affected by each VD category (VD_1-VD_5) during the growing season (May-October). Right panels summarise the average area affected (%) by severity category over the growing season. Results are presented for: (a) observed conditions, (b) the integrated prediction model, (c) the Copula-cp40 model, (d) the EM-f2-0.1 model, and (e) the persistence prediction (PERS) benchmark.

5.5. Discussion

5.5.1. Event-matching based predictive capabilities across land cover types

This section synthesises and expands upon the methodological uncertainties previously discussed in Chapters 3 and 4, with a focus on how they affect the predictive framework developed in this chapter. While many limitations stem from data characteristics and modelling assumptions, I also identify several areas where further refinement is needed to enhance operational applicability.

The analysis shows that predictive skill under the event-matching framework is maximised when SMA and VIA thresholds are closely aligned. This pattern is consistent with results obtained using in-situ soil moisture observations in the United States (Chapter 4; Y. Li *et al.,* 2023), suggesting that threshold alignment represents a robust empirical condition for identifying drought-related vegetation impacts across contrasting data sources. The variability in threshold pattern (Fig 5.5) indicates a non-linear sensitivity of predictive skill to paired SMA and VIA thresholds, consistent with previous studies reporting non-linear vegetation responses to increasing drought stress (Skelton *et al.,* 2015; Vicente-Serrano *et al.,* 2013).

Across southeastern Australia, SMA trigger thresholds derived from the event-matching framework vary substantially by land-cover type and VD severity (Fig. S5.4 and Fig. S5.5). This variability indicates that the level of antecedent soil moisture deficit required to precede vegetation deterioration is not uniform across ecosystems. Agricultural land uses exhibit stronger predictive performance than natural vegetation, especially for moderate to severe VD categories (VD_3–VD_5). The event-matching–based models achieve higher hit rates in cropping and pasture systems than in forested landscapes, reflecting clearer and more temporally coherent vegetation responses to soil moisture deficits in managed systems.

Differences among pasture types are also evident. Model performance is generally stronger for grazing on modified pasture than for grazing on native vegetation, particularly in sub-humid regions. This contrast suggests that vegetation structure and management practices influence the detectability of drought signals in vegetation indices, consistent with findings from previous regional and global studies (X. Li *et al.,* 2023; Guo *et al.,* 2023; Kannenberg *et al.,* 2022; Olson *et al.,* 2018).

The event-matching framework identifies lead times ranging from 0 to 56 days across the study region, with median lead times progressively increasing with VD severity (from 0 days for VD_1 to 32 days for VD_5). This pattern broadly agrees with lead times of 16–48 days reported for in-situ soil moisture–based analyses in the United States (Chapter 4). In contrast, the copula-based approach (Chapter 3) identified consistent lead times across all VD conditions within each pixel, with the temporal scope mainly ranging from 16 to 48 days over the study region. These temporal patterns correspond with global dryland vegetation analyses (Chen *et al.,* 2014; Liu *et al.,* 2023; Tian *et al.,* 2019), indicating shorter response times for shallow-rooted biomes (<1 month) compared to deep-rooted systems (1-2 months). Forest ecosystems present a notable paradox, exhibiting the longest lead times yet limited forecast potential—attributable to their access to deeper soil moisture reserves and complex drought response mechanisms.

5.5.2. Merits and limitations of the integrated predictive models

The integration of land surface model-simulated soil moisture anomalies (SMA) with satellite-derived vegetation index anomalies (VIA) represents a substantive step forward in forecasting vegetation conditions under drought stress. Although none of the proposed models (the event-matching models, the copula-based models, nor the integrated framework) consistently outperforms the persistence (PERS) benchmark across all vegetation deterioration (VD) categories, each demonstrates distinct strengths in capturing different levels of vegetation stress. Importantly, PERS serves here solely as a diagnostic benchmark rather than a viable early warning system, as it does not provide advance warning of drought impacts and con only indicate persistence once vegetation deterioration has already occurred. The performance of the proposed models should therefore be interpreted in light of their capacity to provide anticipatory information, despite the inherent complexity and uncertainty of soil moisture-vegetation interactions, which contracts with the vegetation memory and autocorrelation effects leveraged by PERS. The relatively strong performance of the persistence benchmark is consistent with prior studies using single-variable autoregressive models, such as linear regression (Zambrano *et al.,* 2018), Gaussian process regression (Barrett *et al.,* 2020), and Bayesian autoregression (Salakpi *et al.,* 2022), which exploit the temporal continuity of vegetation indices for short-term forecasting. In contrast, the proposed methods rely on lagged relationships between hydrological and vegetation indicators, which are often weaker and more uncertain. Rather than viewing this as a limitation, it underscores a valid trade-off: while persistence offers competitive baseline accuracy, it is structurally constrained and cannot integrate external hydro-meteorological forecasts or differentiate among categorical risk levels.

The integration of multiple prediction methodologies has emerged as a promising approach to enhance predictive capacity by exploiting complementary strengths across different modelling paradigms (Hao *et al.,* 2016). Previous studies have shown that improved forecasting performance can arise either from combining multiple information sources or from applying multiple modelling strategies to the same predictor to better characterise uncertainty and regime-dependent behaviour (Coelho *et al.,* 2006; Doblas-Reyes *et al.,* 2013). In this study, model integration is achieved through the latter approach. A single information source (i.e., SMA) is used throughout, while distinct modelling frameworks are applied to extract complementary predictive signals. The copula-based methodology demonstrates greater efficacy in identifying non- to low-risk conditions by emphasising statistical dependence and distributional characteristics, while the EM-based approach provides improved detection of severe deterioration by focusing on specific VD events. The integrated framework therefore combines methodological, rather than informational, diversity to balance predictive performance across VD severity categories.

From a risk management perspective, the category-specific optimisation approach ensures the integrated framework maintains appropriate sensitivity across the full spectrum of vegetation deterioration conditions, effectively reducing false alarms in detecting mild vegetation stress while minimising missed detections of severe deterioration events. Underestimation poses greater risks by potentially missing warnings for severe vegetation decline. On the other hand, overestimation, leading to unnecessary alerts, provides a more conservative management approach for the potential heavy economic loss. My category-specific optimisation strategy ensures appropriate sensitivity across the spectrum of vegetation deterioration conditions while reducing false alarms for mild stress and minimising missed detections of severe events.

The multi-categorical prediction framework offers substantial utility for drought early warning systems (DEWS) by establishing hierarchical warning triggers, defining threshold levels that signal escalating risk and inform preparedness or mitigation actions (WMO 2015; Van Loon *et al.,* 2023). The pixel-wise spatial predictions enable the transformation of localised warning signals into regionalised risk assessments by quantifying the areal extent with varying severity categories. This spatial aggregation methodology proffers essential decision support for state and federal drought policy implementation and resource allocation. Through the integration of regionalised risk forecasts with economic indicators, the methodology enables detailed analytical insights into potential agricultural and economic impacts across affected regions. The operational value of these predictions warrants further assessment through measurable economic and strategic benefits realised by stakeholders, as conceptualised in Murphy's (1993) framework for forecast evaluation.

5.5.3. Future study directions

5.5.3.1 Evaluating the added predictive value of soil moisture

A key question emerging from these findings is whether, given the strong temporal autocorrelation in vegetation indices, soil moisture provides any additional predictive skill for vegetation condition. While the persistence benchmark performs strongly in short-term forecasting, it offers no insight into the external drivers of change. In contrast, root-zone soil moisture captures hydrological stress accumulation and can, in principle, provide early warning of vegetation deterioration before it becomes visible in NDVI anomalies. This distinction may be particularly relevant for anticipating turning points in vegetation trajectories, such as the onset of decline following sustained moisture deficits, or for extending forecast lead times beyond the memory range of vegetation itself.

Future research could explore hybrid models that explicitly combine VD persistence with soil moisture anomalies to test whether integrated signals improve predictive skill, particularly for transitions between risk categories. Techniques such as residual skill analysis or mutual information decomposition could help quantify the marginal contribution of soil moisture predictors, after accounting for autocorrelation in VD short-term memories. Answering this question is not only methodologically important but also critical for understanding the value of hydro-meteorological forecasts in operational vegetation impact prediction.

5.5.3.2. Extension of lead time through integrating hydrological forecasts

The predictive capability demonstrated in this study achieves lead times of 2-6 weeks for cropping and grazing land use areas, which provides limited scope for implementing comprehensive drought mitigation strategies. This relatively short prediction window primarily reflects the "propagation time" between soil moisture deficits and vegetation response, as the analysis utilises near-real-time soil moisture estimates rather forecast products with additional lead time.

There is considerable potential to extend these lead times by incorporating skilled hydro-climatological forecasts extending to seasonal timescales. For example, the Australian Water Outlook's seasonal ensemble soil moisture hindcasts demonstrate skill of up to three months ahead (Vogel *et al.,* 2021), suggesting the feasibility of integrating longer-range signals into vegetation impact prediction frameworks. Such extensions could enhance anticipatory capabilities beyond the temporal scope achievable with current observational inputs. .

While this hydro-climatological pathway offers advantages in terms of foresight, the strong performance of persistence-based VD predictions observed in this study suggests that future frameworks should consider incorporating VD persistence components. Persistence captures valuable memory effects in vegetation dynamics, particularly at short lead times, and may serve as a complementary input to the uncertainty-prone nature of forecast-based information. The development of hybrid models that integrate both memory-driven continuity and forward-looking hydro-meteorological anomalies has the potential to produce more robust and balanced predictions. However, such approaches will also need to address the challenges posed by limited sample sizes, especially when modelling extreme events across multiple lead times and land use types.

5.5.3.3. Enhancement of model integration techniques

My current integrated framework demonstrates the complementary strengths of EM-based and copula-based approaches. However, several methodological refinements could further optimise prediction capability. The framework would benefit from exploring more nuanced F-score improvement requirements, particularly in the range of 0.02-0.12 at 0.02 intervals, so as to better balance sensitivity and specificity in drought vegetation impact detection.

The varying false alarm and missing event trade-offs suggest potential for targeted optimisation within specific key subregions, even under the same land use classification. For example, the dryland cropping areas in northeastern New South Wales and western Victoria exhibit distinct performance metric characteristics, despite sharing the same broad land use category. These regional variations indicate opportunities for developing tailored optimisation model settings that account for dominant vegetation species and local climate patterns. Furthermore, implementing more sophisticated assessment techniques, including cross-validation strategies across multiple drought episodes and incorporating economic impact metrics, would enhance the statistical robustness of drought prediction capabilities.

5.6. Conclusions

This study has evaluated the predictive capability of model-simulated root-zone soil moisture for anticipating vegetation drought impacts across southeastern Australia. Three modelling strategies were compared: the threshold-based event-matching (EM) framework, the copula-based probabilistic model, and an integrated EM-copula approach for predicting vegetation deterioration (VD) risks across multiple severity levels. The main findings are summarised as follows:

- The integrated approach improved overall prediction skill by combining the strengths of its components. The copula-based approach excelled in detecting mild vegetation deterioration stress (VD_{1-2}) with lower false alarm rates. The EM method provided more balanced predictions for severe deterioration risks (VD_{3-5}), achieving hit rates between 0.25 and 0.58 for extreme categories whilst maintaining acceptable adjacent category accuracy.

- Agricultural lands showed higher predictive capability compared to non-agricultural regions, with lead times varying systematically by vegetation type. Prediction windows ranged from 0-32 days in agricultural areas to 16-56 days in forested regions, reflecting the influence of vegetation rooting depth and local climate patterns on drought response timing. Grazing modified pastures exhibited particular sensitivity to soil moisture deficits, whereas forest ecosystems showed longer but less predictable response patterns.

- Stricter SMA trigger selection criteria led to substantially more conservative VD predictions. In the copula-based framework, increasing the exceedance probability threshold from 40% to 70% reduced false alarms but decreased the spatial coverage of valid triggers by more than 20%. Similarly, in the event-matching framework, raising the F-score improvement criterion from 0.1 to 0.3 reduced trigger coverage by up to one third, quantifying a clear trade-off between predictive conservatism and geographic applicability.

- Temporal aggregation from 8-day to monthly scales significantly improved predictive performance. Hit rates for severe deterioration increased from 0.16 to 0.29, with reduced underestimation errors and stable adjacent prediction ratios. Monthly-scale aggregation enhanced detection of sustained vegetation stress by filtering short-term variability, although this comes at the cost of finer temporal resolution.

While the integrated and component models demonstrated context-specific strengths, none consistently outperformed the persistence benchmark. This highlights the strong temporal autocorrelation in vegetation index anomalies (VIA), which supports the effectiveness of persistence-based approaches in short-term forecasting. Rather than viewing this as a limitation, future research should consider incorporating vegetation condition persistence as a predictor within the modelling framework. Such hybrid approaches, combining memory-based signals with hydro-meteorological drivers, may offer enhanced skill in anticipating the onset, persistence, and recovery of vegetation stress under drought.

Appendix A

Appendix Text A1.

To demonstrate my verification metrics framework, I present a detailed computational analysis using the severe vegetation deterioration category (VD_3) as extracted from the multi-category contingency table. This analysis specifically examines the prediction outcomes from the EM-f2-0.1 set of Soil Moisture Anomaly (SMA) trigger thresholds (Figure S1). The verification metrics are calculated as follows:

1. Hit Rate (Probability of Detection)

- hits = 58,359 (diagonal element)
- misses = 163,401 + 29,046 + 48,994 + 33,576 + 20,327 = 295,344 (sum of off-diagonal row elements)

Hit Rate = hits / (hits + misses) =58,359 / (58,359 + 295,344) ≈ 0.16

2. False Alarm Ratio

- false alarms = 181,373 + 97,825 + 97,886 + 12,978 + 3,911 = 393,973 (sum of off-diagonal column elements)

FAR = false alarms / (hits + false alarms) = 393,973 / (58,359 + 393,973) ≈ 0.87

3. Underestimation Ratio

- Underestimated events = 163,401 + 29,046 + 48,994 = 241,441 (sum of row elements where predicted VD categories < true category VD3)
- Total observations in VD 3 = 353,703 (sum of row elements of true category 3)

Underestimation ratio = 241,441 / 353,703 × 100% ≈ 68%

4. Overestimation Ratio

- Overestimated events = 33,576 + 20,327 = 53,903 (sum of row elements where predicted VD categories > true category VD3)

Overestimation ratio = 53,903 / 353,703 × 100% ≈ 15%

5. Adjacent Prediction Ratio

- Adjacent predictions = 48,994 + 33,576 = 82,570 (Predicted VD category in true category of VD2 and VD4)
- Total misclassifications = 295,344 (all non-diagonal elements in row VD3)

APR = (82,570 / 295,344) × 100% ≈ 28%

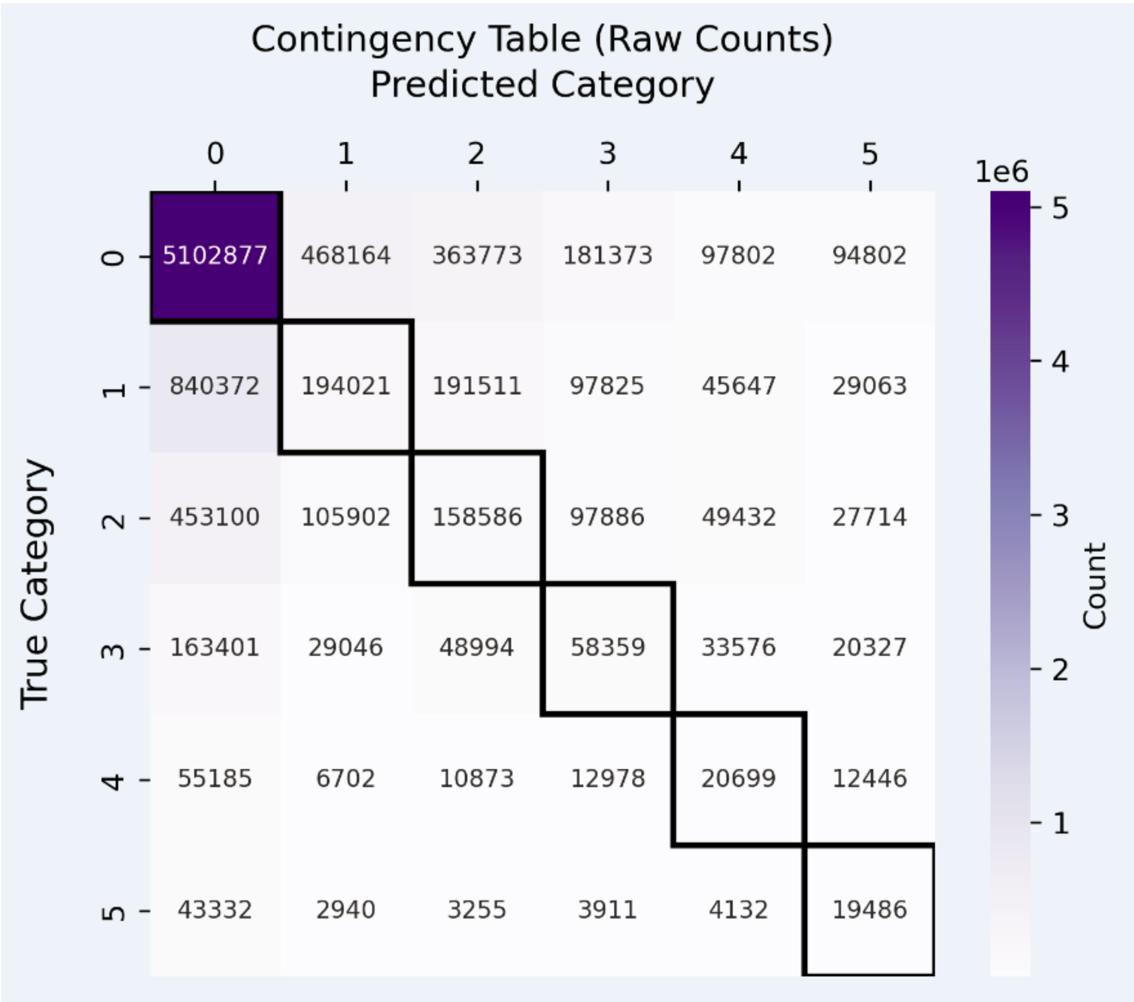


Fig. T1A1. Multi-category contingency table illustrating the distribution of predicted versus observed (true) categorical vegetation deterioration (VD) events. The matrix displays raw counts of classification outcomes across six VD categories (0-5). Diagonal elements (highlighted with boxes) represent correct predictions, whilst off-diagonal elements indicate misclassification frequencies. The colour gradient from light to dark purple corresponds to increasing frequency counts, with the scale indicating values in millions (×10^6).

Supplementary Information for Chapter 5

SI Figures

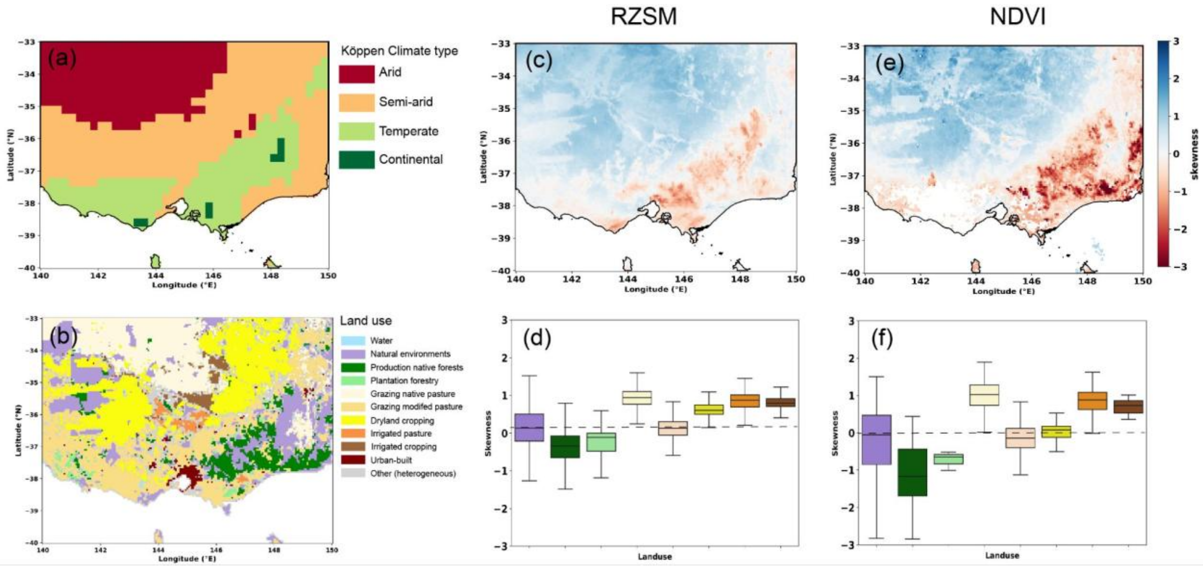


Fig. S5.1. Spatial distribution of (a) Climate zones classified according to the Köppen-Geiger system (Beck *et al.,* 2018), and (b) land use types. Panels (c) and (e) map the skewness of Root Zone Soil Moisture (RZSM) and Normalized Difference Vegetation Index (NDVI), respectively, from 2001 to 2022. Box plots (d) and (f) display the distribution of skewness for RZSM and NDVI across different land use categories. Areas depicted as blank grids represent aquatic bodies.

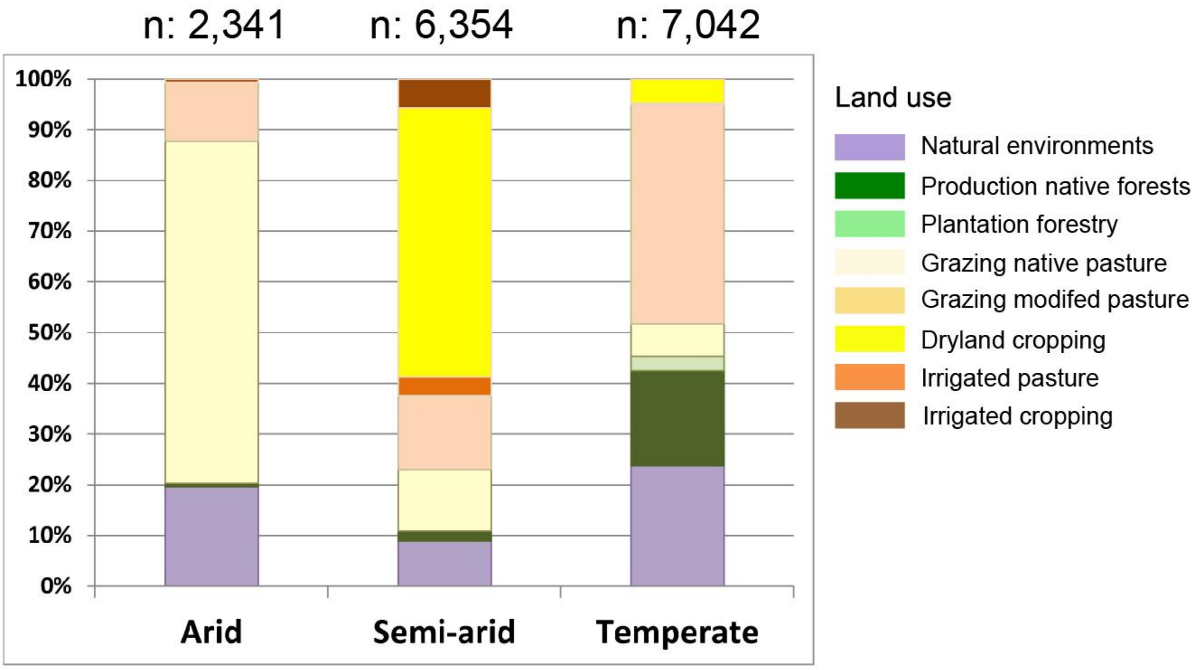


Fig. S5.2. Percent stacked bar chart shows the prevalence of various land use types within arid, semi-arid, and temperate climates across the study area. The number of grids sampled for each climate zone is denoted by the ‘n’ value above the respective bars. In summary, grazing native pastures are primarily situated in arid zones, dryland cropping dominates semi-arid regions, and modified pastures are the main land use in temperate climates.

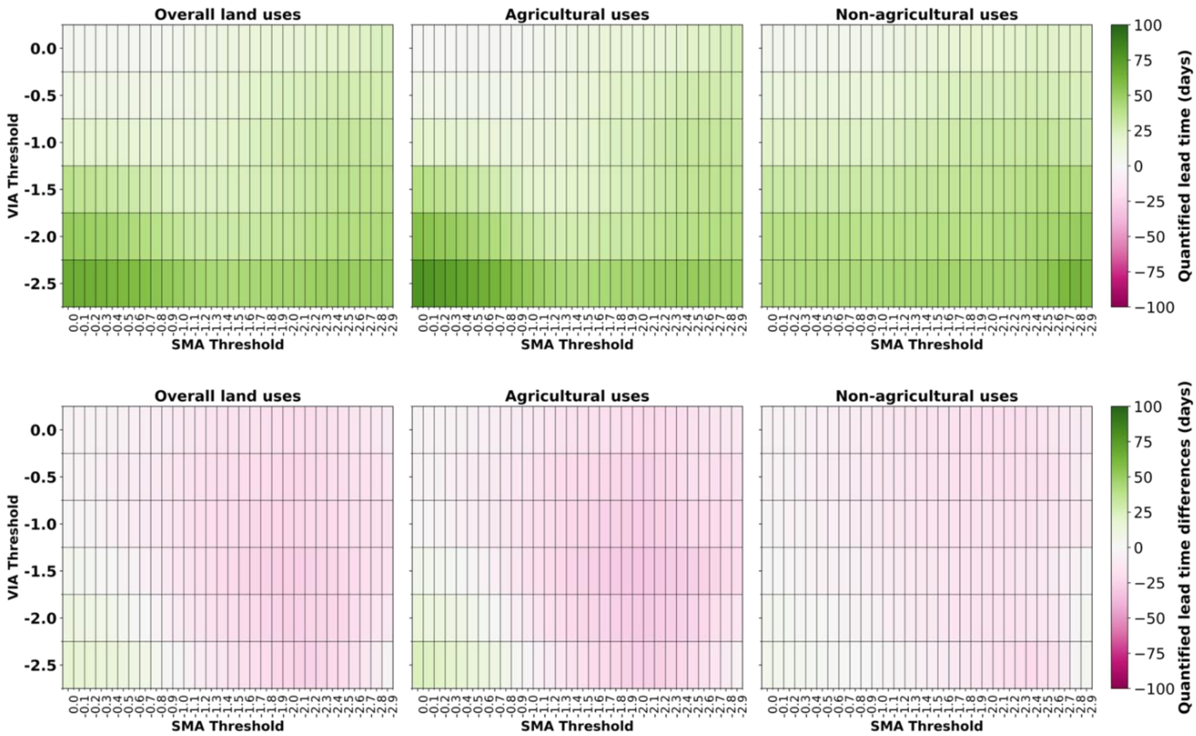


Fig. S5.3. Quantified *Lead times* (top row) and *Lead time* differences (bottom row) in F_{β}-score for each combination of vegetation index anomaly (VIA) and soil moisture anomaly (SMA) thresholds, relative to a bootstrapped random baseline. Results are summarised by land use types: overall, agricultural and non-agricultural. For baseline generation methodology, see Section 5.3.1.3.

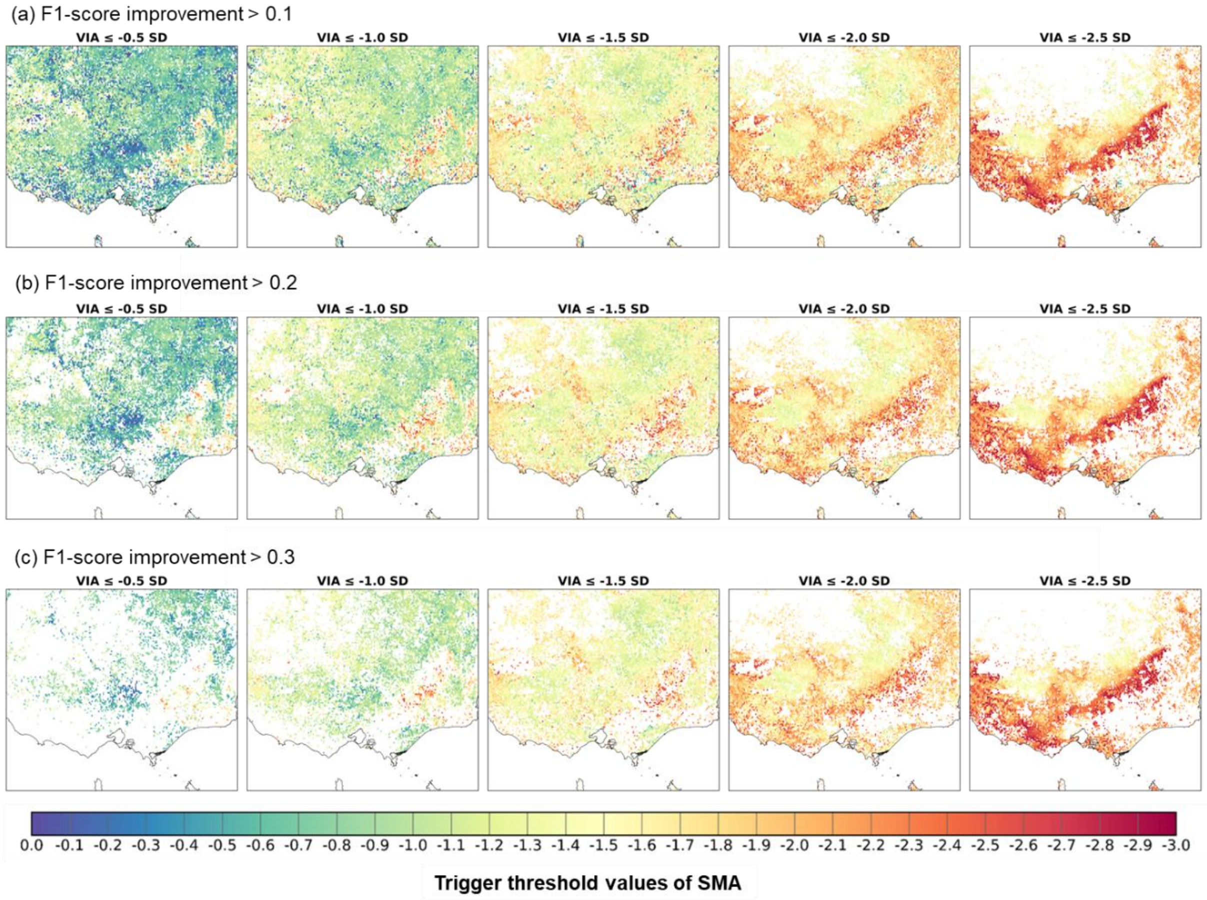


Fig. S5.4. Spatial distribution of soil moisture anomaly (SMA) trigger thresholds for different vegetation deterioration (VD) categories, identified using F_1-score improvements above (a) 0.1, (b) 0.2, and (c) 0.3. Each row (a–c) shows SMA trigger thresholds corresponding to increasing VD severity levels: mild (VIA ≤ -0.5 SD), moderate (VIA ≤ -1.0 SD), severe (VIA ≤ -1.5 SD), extreme (VIA ≤ -2.0 SD), and exceptional (VIA ≤ -2.5 SD).

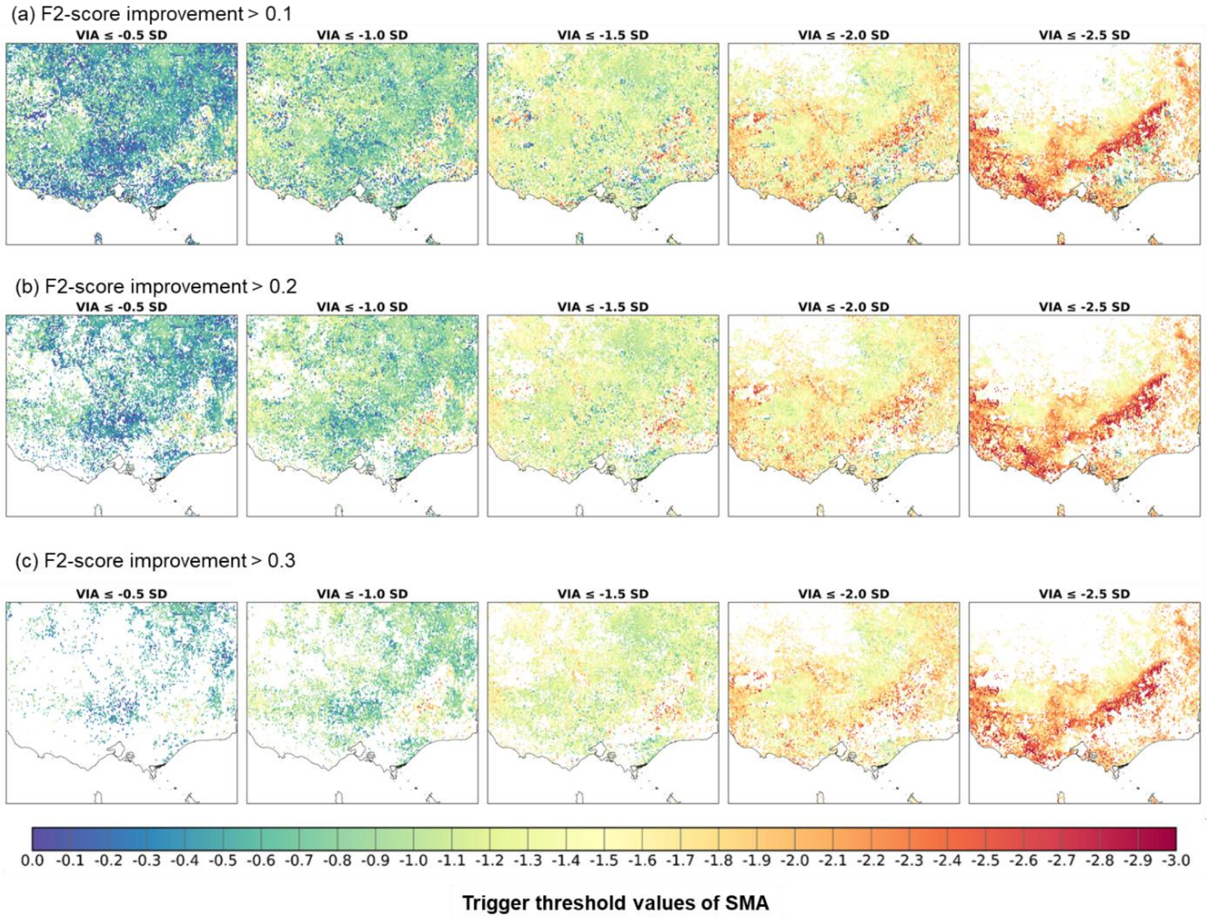


Fig. S5.5. same for fig. S5.4 but for F_2-score improvements.

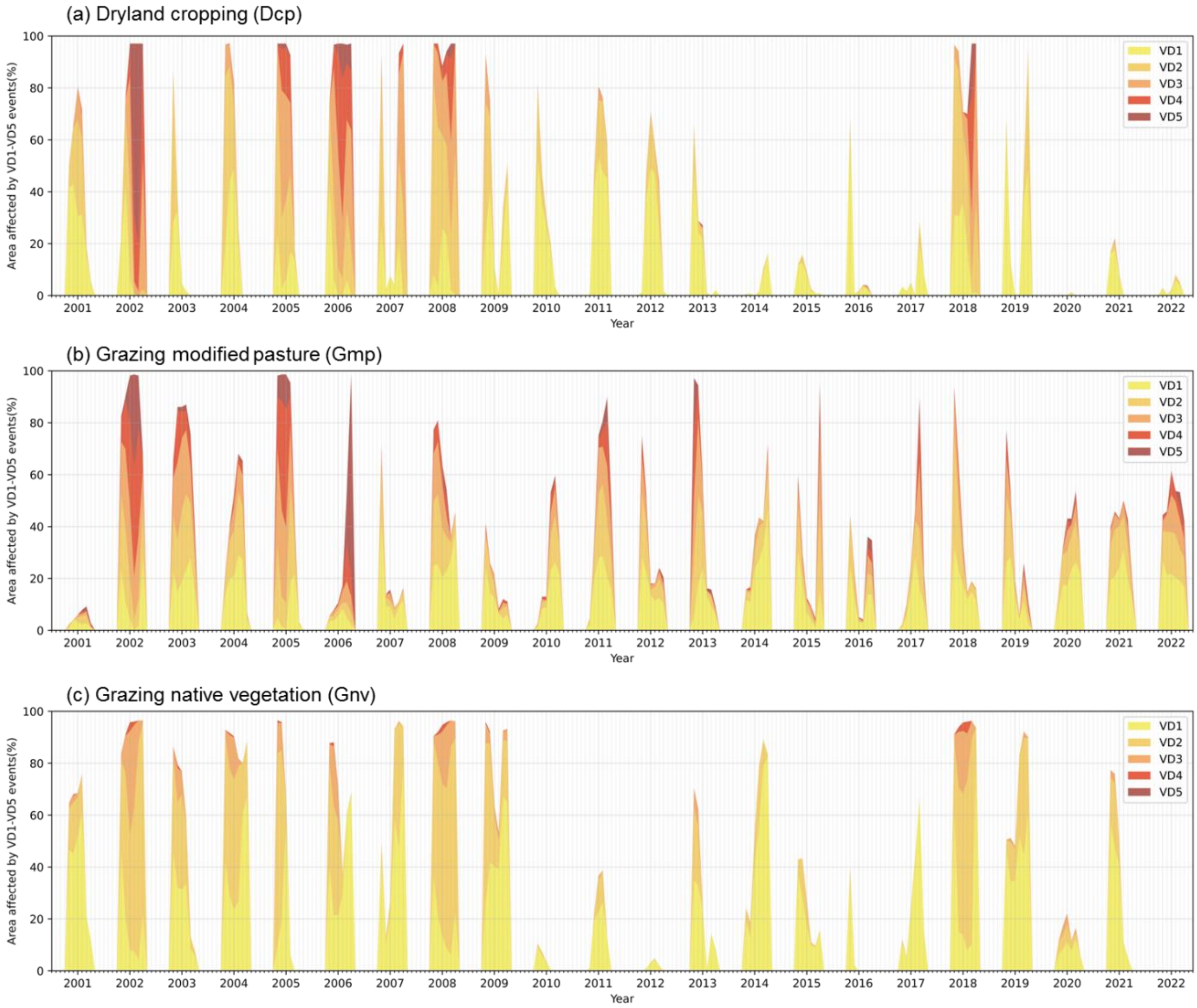


Fig. S5.6. Temporal patterns of vegetation deterioration across three key agricultural subregions during 2001-2022. The stacked area plots represent the area under vegetation deterioration condition (VD_1-VD_5 categories) during each growing season of 2001-2022 for (a) dryland cropping region, (b) grazing modified pasture, and (c) grazing native vegetation. The severity of vegetation deterioration increases from VD_1 (mild, yellow) to VD_5 (exceptional, dark red), with the total height indicating the cumulative affected area as a percentage of each subregion.

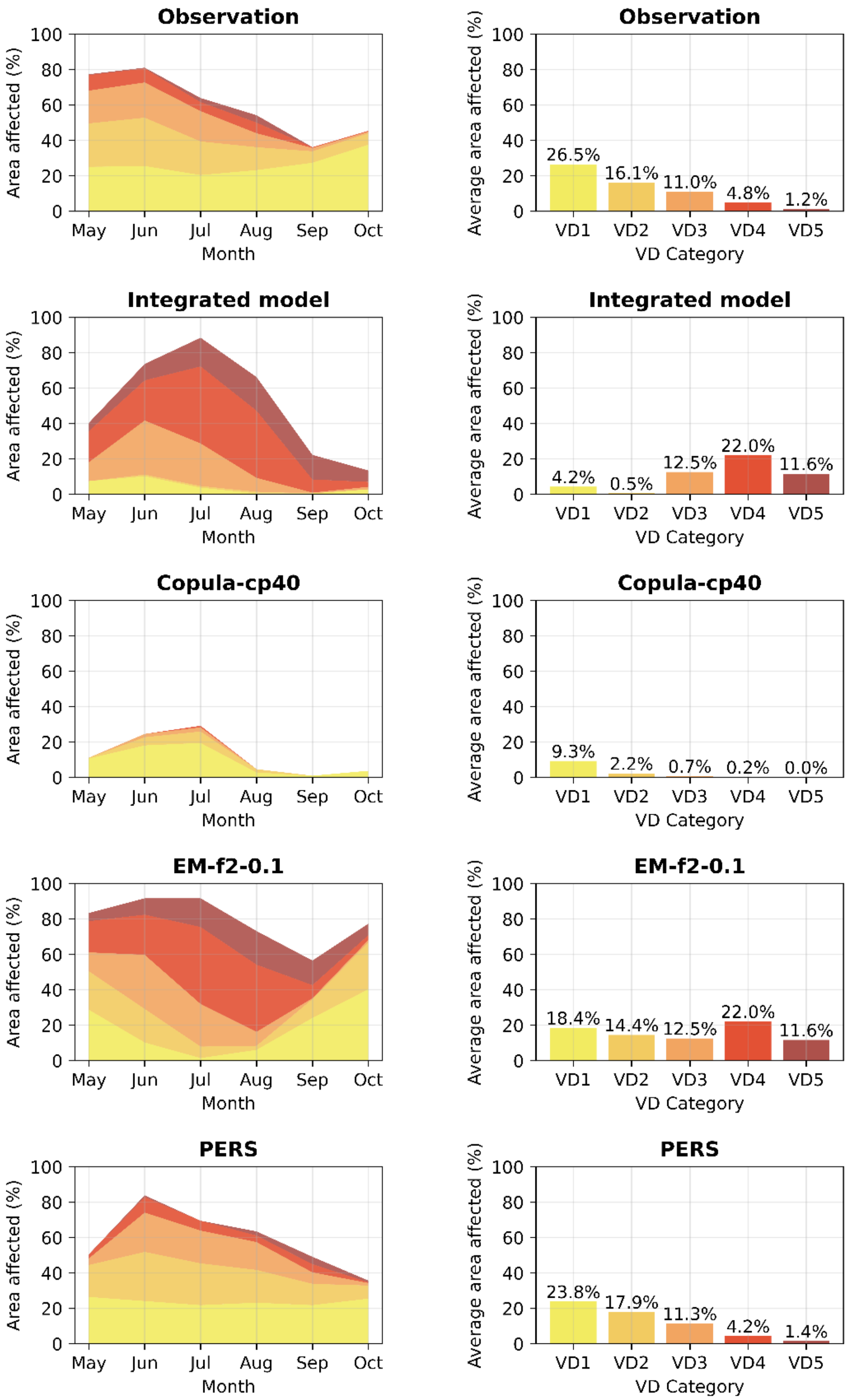


Fig. S5.7. Same as Fig. 5.8, but for the grazing modified pasture (Gmp) case study region during the 2008 drought year.

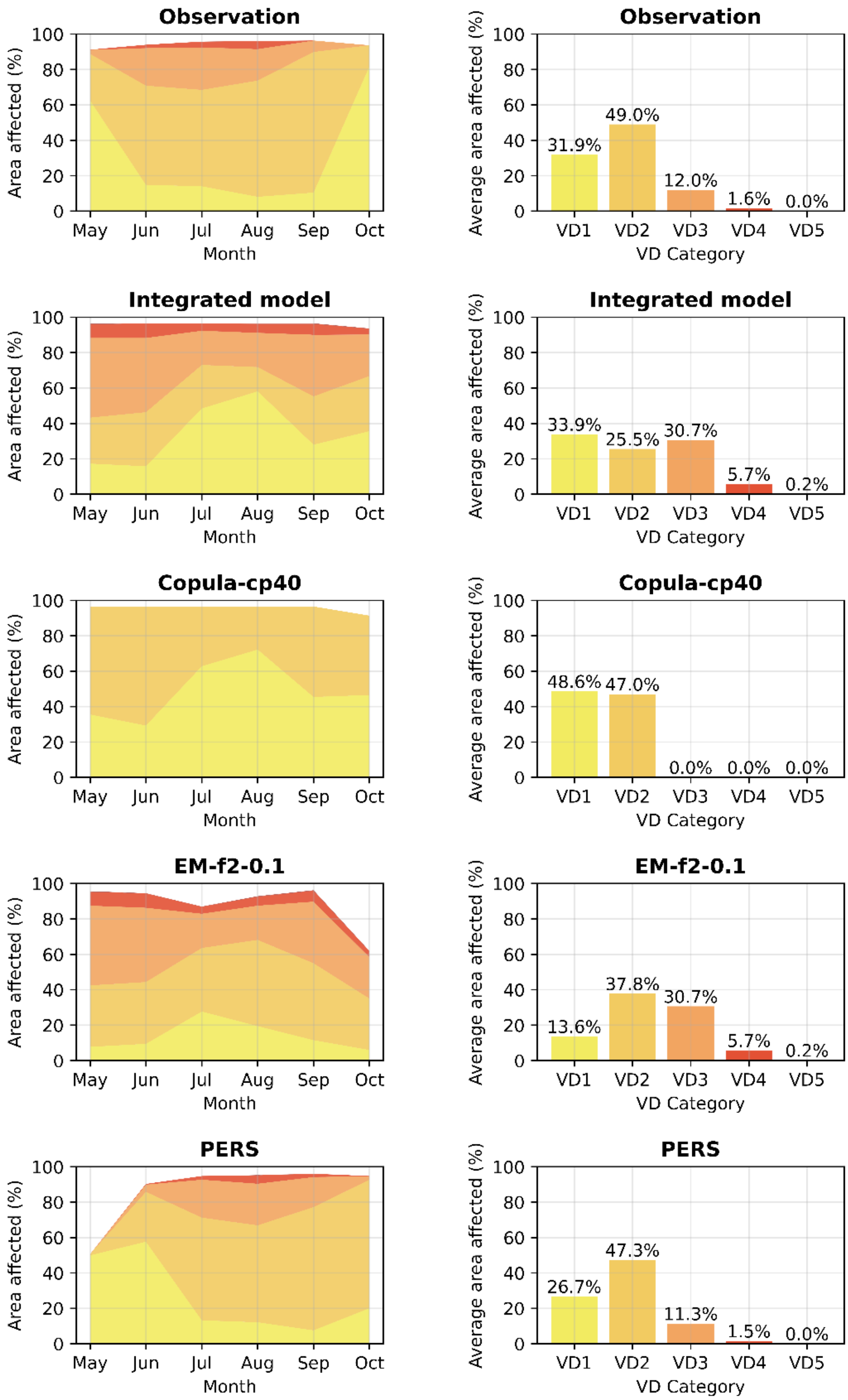


Fig. S5.8. Same as Fig. 5.8, but for the grazing native vegetation (Gnv) case study region during the 2018 drought year.

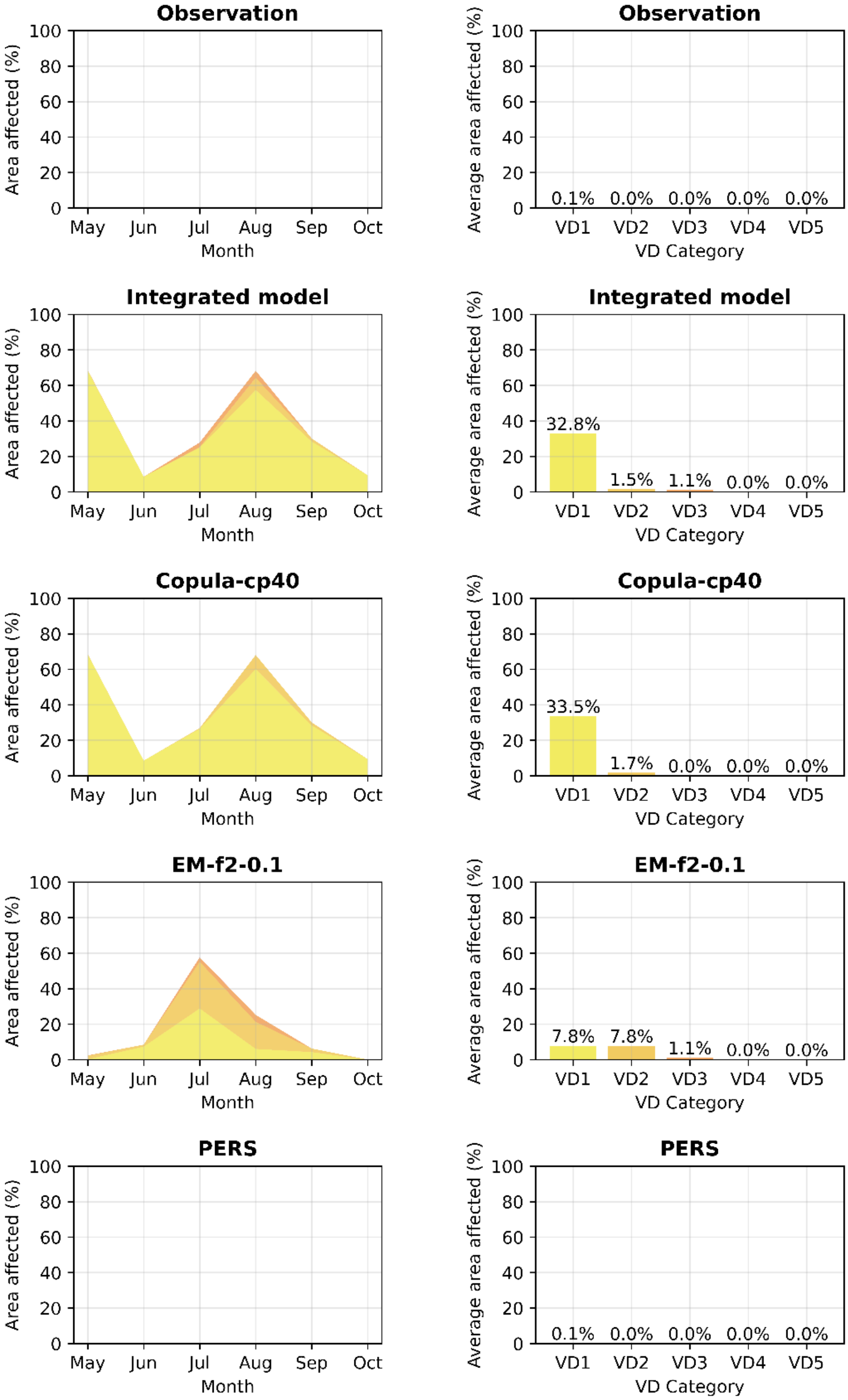


Fig. S5.9. Same as Fig. 5.8, but for the grazing native vegetation (Gnv) case study region during the 2022 non-drought year.

Supplementary Tables

Table S5.1. Median *Lead time* (days) of vegetation deterioration (VD) risk prediction using multiple estimated SMA trigger sets.

VD risk category	EM-f1-0.1	EM-f1-0.2	EM-f1-0.3	EM-f2-0.1	EM-f2-0.2	EM-f2-0.3
VD_1	0	0	0	0	0	0
VD_2	8	8	8	8	8	8
VD_3	16	16	16	16	16	16
VD_4	24	24	24	24	24	24
VD_5	32	32	40	32	32	32

Table S5.2. Summary statistics of prediction performance metrics for each SMA sets at an 8-day temporal scale.

SMA sets name	VD Severity	Hit rate	FAR	Under-estimation (%)	Over-estimation (%)	APR (%)
Copula-cp40	0	0.87	0.22	0	13	88
	1	0.33	0.73	60	7	99
	2	0.14	0.71	85	1	44
	3	0.02	0.74	97	1	21
	4	0.02	0.79	98	0	5
	5	0.01	0.84	99	0	2
	overall	0.23	0.67	73	4	43
Copula-cp50	0	0.94	0.25	0	6	91
	1	0.21	0.71	77	3	100
	2	0.06	0.7	94	1	30
	3	0.01	0.73	98	0	11
	4	0.01	0.78	99	0	3
	5	0	0.86	100	0	1
	overall	0.21	0.67	78	2	39
Copula-cp60	0	0.97	0.28	0	3	93
	1	0.11	0.72	88	1	100
	2	0.02	0.72	98	0	19
	3	0.01	0.71	99	0	5

	4	0.01	0.77	99	0	2
	5	0	0.89	100	0	1
	overall	0.19	0.68	81	1	37
Copula-cp70	0	0.99	0.29	0	1	93
	1	0.05	0.75	95	0	100
	2	0.01	0.75	99	0	10
	3	0.01	0.69	99	0	2
	4	0	0.74	99	0	1
	5	0	0.89	100	0	0
	overall	0.18	0.69	82	0	34
EM-f1-0.1	0	0.86	0.24	0	14	49
	1	0.15	0.74	68	17	92
	2	0.15	0.77	72	13	25
	3	0.13	0.84	77	10	23
	4	0.14	0.89	79	7	20
	5	0.21	0.86	79	0	5
	overall	0.27	0.72	63	10	36
EM-f1-0.2	0	0.91	0.26	0	9	49
	1	0.11	0.72	76	13	94
	2	0.13	0.74	77	10	19
	3	0.12	0.81	80	8	20
	4	0.12	0.87	82	6	17
	5	0.18	0.81	82	0	4
	overall	0.26	0.7	66	8	34
EM-f1-0.3	0	0.96	0.28	0	4	34
	1	0.04	0.71	88	8	96
	2	0.08	0.72	84	8	10
	3	0.1	0.78	83	7	14
	4	0.11	0.85	84	5	15
	5	0.16	0.76	84	0	4
	overall	0.24	0.68	71	5	29
EM-f2-0.1	0	0.81	0.23	0	19	39
	1	0.14	0.76	60	26	86
	2	0.18	0.8	63	20	28
	3	0.16	0.87	68	15	28
	4	0.17	0.92	72	10	26

	5	0.25	0.9	75	0	7
	overall	0.29	0.75	56	15	36
EM-f2-0.2	0	0.87	0.25	0	13	38
	1	0.1	0.74	69	21	90
	2	0.15	0.78	68	17	21
	3	0.15	0.85	72	13	24
	4	0.16	0.9	75	9	23
	5	0.22	0.87	78	0	6
	overall	0.28	0.73	60	12	34
EM-f2-0.3	0	0.95	0.27	0	5	23
	1	0.03	0.73	84	13	93
	2	0.09	0.75	78	13	11
	3	0.13	0.82	76	11	17
	4	0.14	0.88	78	8	19
	5	0.19	0.84	81	0	5
	overall	0.26	0.72	66	8	28
PERS	0	0.88	0.12	0	12	73
	1	0.45	0.58	38	17	94
	2	0.41	0.57	47	12	65
	3	0.33	0.62	58	9	55
	4	0.29	0.66	63	8	44
	5	0.52	0.39	48	0	27
	overall	0.48	0.49	42	10	60

Table S5.3. Summary statistics of prediction performance metrics for each SMA sets at a monthly aggregated temporal scale.

SMA sets name	VD Severity	Hit rate	FAR	Under-estimation (%)	Over-estimation (%)	APR (%)
Copula-cp40	0	0.75	0.26	0	25	86
	1	0.46	0.72	42	13	99
	2	0.22	0.69	76	1	61
	3	0.04	0.72	95	1	33
	4	0.03	0.73	96	1	8
	5	0.03	0.7	97	0	5
	overall	0.25	0.64	68	7	49
Copula-cp50	0	0.87	0.32	0	13	92
	1	0.33	0.72	62	5	100
	2	0.09	0.7	90	1	45
	3	0.02	0.71	97	1	18
	4	0.02	0.72	98	1	5
	5	0.02	0.7	98	0	3
	overall	0.23	0.64	74	3	44
Copula-cp60	0	0.94	0.36	0	6	94
	1	0.19	0.73	79	1	100
	2	0.03	0.73	96	1	30
	3	0.01	0.7	98	1	8
	4	0.01	0.7	98	0	3
	5	0.01	0.72	99	0	2
	overall	0.2	0.66	79	1	40
Copula-cp70	0	0.98	0.39	0	2	94
	1	0.09	0.76	91	0	100
	2	0.01	0.77	99	0	16
	3	0.01	0.69	99	0	3
	4	0.01	0.65	99	0	2
	5	0.01	0.77	99	0	2
	overall	0.18	0.67	81	1	36
EM-f1-0.1	0	0.77	0.26	0	23	51
	1	0.23	0.71	48	29	84
	2	0.25	0.71	53	22	43

	3	0.24	0.77	59	17	41
	4	0.28	0.81	61	11	38
	5	0.52	0.77	48	0	18
	overall	0.38	0.67	45	17	46
EM-f1-0.2	0	0.84	0.3	0	16	49
	1	0.18	0.7	60	23	88
	2	0.22	0.69	60	18	34
	3	0.23	0.74	63	14	36
	4	0.26	0.79	64	9	35
	5	0.48	0.74	52	0	18
	overall	0.37	0.66	50	13	43
EM-f1-0.3	0	0.93	0.35	0	7	33
	1	0.06	0.69	79	15	92
	2	0.14	0.67	72	14	17
	3	0.2	0.72	69	11	25
	4	0.24	0.77	68	8	29
	5	0.43	0.7	57	0	17
	overall	0.33	0.65	58	9	36
EM-f2-0.1	0	0.69	0.24	0	31	42
	1	0.21	0.73	38	42	75
	2	0.28	0.75	41	32	47
	3	0.29	0.81	47	24	48
	4	0.34	0.85	51	15	48
	5	0.58	0.81	42	0	24
	overall	0.4	0.7	37	24	47
EM-f2-0.2	0	0.77	0.28	0	23	38
	1	0.15	0.72	50	35	80
	2	0.24	0.73	48	28	37
	3	0.28	0.79	51	21	42
	4	0.33	0.83	54	13	45
	5	0.56	0.79	44	0	24
	overall	0.39	0.69	41	20	44
EM-f2-0.3	0	0.89	0.34	0	11	22
	1	0.04	0.71	73	23	87
	2	0.14	0.71	63	23	19
	3	0.24	0.77	58	18	30

	4	0.31	0.82	58	12	38
	5	0.53	0.77	47	0	22
	overall	0.36	0.68	50	14	36
PERS	0	0.84	0.14	0	16	71
	1	0.52	0.53	29	18	93
	2	0.51	0.5	38	11	71
	3	0.45	0.5	46	8	62
	4	0.43	0.5	51	5	50
	5	0.48	0.39	52	0	31
	overall	0.54	0.43	36	10	63

Table S5.4. Performance requirements for vegetation deterioration (VD) predictability and best-performing SMA trigger threshold sets at monthly aggregated temporal scale.

Best-performance Requirements	Best set of SMA trigger thresholds	Best score from the model	Perfect score
1: High hit rate for severe deterioration (VD_3-VD_5)	EM-f2-0.1	VD_3: 0.29, VD_4: 0.34, VD_5: 0.58	1.0
2: Minimal underestimation in VD_4-VD_5 with high APR	EM-f2-0.1	VD_4: 51% (48%), VD_5: 42% (24%)	0% (100%)
3: Low FAR in VD_1-VD_2 predictions	EM-f1-0.3	VD_1: 0.69 VD_2: 0.67	0.0
4: Minimal overestimation in VD_0-VD_2 with high APR	Copula-cp70	VD_0: 2% (94%) VD_1: 0% (100%) VD_2: 0% (16%)	0% (100%)
5: High total APR	Copula-cp40	Overall: 49%	100%

Note: Values in parentheses denote the complementary misclassification metric associated with the primary performance indicator (e.g. underestimation or adjacent prediction ratio, APR).

Chapter 6

Probabilistic early warning of drought-induced vegetation deterioration risks: using seasonal soil moisture forecasts for agricultural regions of southeastern Australia

Abstract

Proactive and effective drought risk management in agricultural regions requires early warning systems capable of accurately forecasting drought-induced vegetation deterioration (VD). This study evaluates the performance of multi-categorical VD forecasts derived from ensemble root-zone soil moisture forecasts at lead times ranging from one to six months, covering agricultural regions in southeastern Australia between 2001 and 2017. VD events were identified from satellite-derived vegetation index anomalies. The study assesses the probabilistic and deterministic skill, accuracy and reliability of these forecasts, comparing them to observations and climatology at a monthly temporal resolution. In particular, I focus on the performance of the Australian Water Outlook's (AWO) soil moisture hindcasts and historical simulations in predicting VD categories, with a specific emphasis on selected key grazing and cropping zones. The results indicate the practical early warning horizons for most regions, spanning two to four months for mild to severe VDs categories. This horizon combines the inherent vegetation-soil moisture response time (0-2 months) and the extended lead time (1-3 months) offered by soil moisture forecasts. While forecast accuracy decreases with longer lead times, the ensemble-based methodology provides valuable probabilistic insights into likelihood and severity of drought-induced vegetation stress. Deterministic forecasts, based on ensemble means, maintain accuracy up to a two-month lead time but tend to underestimate VD risks at longer lead times. The framework exhibits enhanced predictive performance during the late growing season (August–October) compared to the early season (May–July), particularly in inland pastoral and dryland cropping zones. A validation using the 2002 drought case study confirms the framework's operational efficacy, achieving over 50% probabilistic early warning confidence for observed vegetation deterioration risks at three-month lead times. This research underscores the operational value of integrating ensemble soil moisture forecasts with impact-based thresholds for early vegetation stress warning across Australian agricultural regions. The framework's ability to translate soil moisture forecasts into actionable multi-categorical vegetation deterioration warnings across diverse spatial scales facilitates targeted interventions at both local and regional decision-making for agricultural sectors.

6.1. Introduction

Drought is a complex and costly natural hazard with profound societal and environmental consequences. Its impacts are particularly pronounced in agricultural sectors, leading to substantial crop and livestock losses (Chawla *et al.,* 2020; Rajsekhar *et al.,* 2015) and affecting ecosystem functioning (Seneviratne *et al.,* 2012; Zscheischler *et al.,* 2018). In Australia, drought significantly influences rural communities, particularly in dryland farming regions. During the Millennium Drought (2001-2009), wheat yields in Victoria dropped by up to 60% (Stephens and Lyons, 2015), while Queensland and New South Wales saw extensive vegetation decline in pasture land and substantial destocking (McKeon *et al.,* 2004). The 2002-2003 drought reduced national agricultural production by nearly 25%, causing reduction of Australia's GDP by 1.6% (ABARES, 2004; Van Dijk *et al.,* 2013).

Drought early warning systems (DEWS) are designed for mitigating these impacts, relying on hydro-meteorological forecasting tools to provide advance information of drought occurrence and development (Van Loon *et al.,* 2016). Existing subseasonal-to-seasonal hydrological forecasting tools focus on precipitation and streamflow at global or continental-scale, including GloFAS, WW-HYPE and EFAS, and are mainly established for Europe and Northern America (Arheimer *et al.,* 2020; Arnal *et al.,* 2018; Donnelly *et al.,* 2016; Emerton *et al.,* 2018; Vogel *et al.,* 2021). Soil moisture is fundamental for vegetation and agricultural productivity, making it a key variable in drought monitoring and forecasting (Seneviratne *et al.,* 2010). However, operational soil moisture forecasting at the continental scale remains limited, as it depends on complex interactions among the climate conditions, hydrological dynamics, and land surface characteristics (e.g., soil properties, vegetation types) (Shen *et al.,* 2012; Sun *et al.,* 2002; Yin *et al.,* 2018). Soil moisture forecasts are generated by coupling atmosphere-ocean-land general circulation models (GCMs) with land surface models (LSMs), where process-based hydrological models are forced using seasonal climate forecasts produced through numerical weather prediction systems (Arnal *et al.,* 2018; Emerton *et al.,* 2018). Forecasting uncertainties arise from the limited validation from sparse spatially-distributed soil moisture observation sites, particularly in the root-zone depth (0-1m) (Yuan *et al.,* 2015), which further affects forecast reliability at high spatial resolutions (e.g., 5km) (Vogel *et al.,* 2021). Climate model outputs often contain systematic biases in hydrological variables, and failing to correct them can lead to unrealistic drought frequency projections, even at continental scales (Johnson & Sharma, 2015).

Existing operational soil moisture forecasting systems include ECMWF's SEAS5 (Johnson *et al.,* 2019; Ferranti and Weightman, 2019), the UK Met Office's GloSea5 (MacLachlan *et al.,* 2014), and NASA's FLDAS-Forecast version 2 (FFV2) incorporating NOAA's North American Multi-Model Ensemble (NMME) (Hazra *et al.,* 2023). However, most of these systems are developed for Europe and North America (Esit *et al.,* 2021; Sehgal and Sridhar, 2019), limiting their application to other regions.

In Australia, the Bureau of Meteorology's Australian Water Outlook (AWO) provides operational seasonal (3-month ahead) root-zone soil moisture forecasts at 5-km resolution (Hudson *et al.,* 2017; Wilson *et al.,* 2022), which is tailored to Australia's regional climate and hydrology characteristics. These systems widely adopt ensemble forecasting approaches, including multi-model ensembles and single-model ensembles with different initial conditions. This method enhances seasonal hydrological forecasting by leveraging multiple simulation members with varying configurations (Hao *et al.,* 2017; AghaKouchak *et al.,* 2022). It is particularly valuable for soil moisture prediction, as it accounts for uncertainties in both atmospheric forcing and land surface processes (Arnal *et al.,* 2018; Emerton *et al.,* 2018; Pechlivanidis *et al.,* 2017).

Despite advances in seasonal hydro-meteorological forecasting, a persistent gap remains in translating numerical forecasts into formally defined, impact-based drought intelligence that can be systematically operationalised within drought early warning systems (DEWS) (Mishra, 2023). Farmers often develop strong, locality-specific understandings of how climatic variables affect vegetation conditions. However, commonly used drought indices such as the standardised soil moisture index (SSMI; AghaKouchak, 2016) are not expressed in terms of explicit impact severity categories. For example, a forecast SSMI value of -2 is statistically meaningful but does not directly indicate whether vegetation stress is expected to be moderate, severe, or extreme, which limits its use in multi-category, impact-based forecasting and formal drought risk communication (Mishra and Singh, 2010, 2011). Research on multi-category, impact-based drought forecasting remains limited, with only a few studies conceptually describing (Boult *et al.,* 2022; Tozier de la Poterie *et al.,* 2023) or developing and testing such forecasts (e.g., Sutanto *et al.,* 2019; Westerveld *et al.,* 2021; AghaKouchak *et al.,* 2022). An effective approach is to define and align drought-impact thresholds that directly link hydro-meteorological conditions (e.g., soil moisture deficits) to categorical responses (e.g., vegetation stress levels) (Von Loon *et al.,* 2016). Identifying appropriate thresholds (commonly termed "alarm" or "trigger" values) enhance the ability of hydro-meteorological forecasts to provide actionable impact insights for stakeholders through DEWS. In my previous research, I developed a framework for predicting multi-categorical vegetation deterioration events (VDs) by establishing trigger thresholds for soil moisture anomalies over large regions. This vegetation drought impact predicting approach integrates copula-based and event-matching approach with a severity-specific focus (Chapter 5) and has demonstrated promising predictive performance across diverse land use types in southeastern Australia, particularly for agricultural sectors.

This chapter advances my research on categorical drought-induced vegetation impact prediction by evaluating the integration of ensemble soil moisture forecasts into the predictive framework established in Chapter 5. The primary objective is to assess the predictive skill and potential skilful lead time extension achieved by using ensemble soil moisture forecasts. This aligns with Research Question 4 (RQ4) in Chapter 1 (Section 1.2), which focuses on the operational utility of probabilistic soil moisture forecasts in predicting vegetation deterioration risks across different land systems.

While previous studies have shown that ensemble hindcasts (retrospective forecasts) generally exhibit lower accuracy than reanalysis-driven historical simulations due to inherent biases and increasing uncertainty at longer lead times (Hudson *et al.,* 2017; BoM, 2019; Vogel *et al.,* 2021), assessing their practical utility remains essential for informing early warning applications. This analysis aims to quantify the trade-off between forecast lead time and prediction accuracy in operational contexts, such as providing actionable intelligence for proactive drought preparedness and mitigation. Furthermore, the ensemble structure allows probabilistic characterisation of forecast uncertainty, providing stakeholders with information not only on the likelihood of drought-induced impacts but also on severity levels (Stockdale *et al.,* 2010).

Specifically, this chapter assesses the deterministic and probabilistic skill of forecasting vegetation deterioration events (VDs) using ensemble soil moisture hindcasts (with 1- to 6-month lead times) across southeastern Australia at monthly temporal scales. The following research questions represent more specific sub-questions under RQ4, addressing critical dimensions of operational implementation:

i) What is the maximum lead time at which soil moisture forecasts maintain skilful predictive capability for VD risks?

ii) How do the predictive skill and achievable lead times vary across different agricultural land use types?

iii) What valuable insights can be drawn from probabilistic and deterministic (i.e., ensemble mean) VD risk forecasts intro to drought early warning systems at regional scale?

This analysis draws on ensemble soil moisture hindcasts for the period 2001–2017, rather than operational real-time forecasts (available only from November 2021) provided by the Australian Water Outlook. The hindcasts offer a consistent retrospective dataset covering multiple drought cycles and thus enable statistically robust validation of forecast skill. Further methodological details on the hindcast and operational systems are available at: https://awo.bom.gov.au/about/overview/forecasts. By rigorously assessing predictive performance and uncertainty across agricultural systems, this chapter informs the operational implementation of seasonal drought impact forecasting. It strengthens the foundation for evidence-based drought preparedness and policy development, advancing early warning capabilities through robust validation metrics and uncertainty quantification frameworks tailored to vegetation stress prediction.

6.2. Data

In this study, a combination of hindcast (retrospective forecast) and near-real-time monitoring (historical simulation) datasets was utilised to evaluate the predictive skill of forecasting drought-induced vegetation deteriorating cross multiple severity levels in southeastern Australia. A summary of dataset resolution, unit and temporal coverage is provided in Table 6.1, with detailed descriptions in subsequent sections. All datasets share a spatial resolution of 0.05° and a common reference period from 2001 to 2017. The daily hindcast and historical soil moisture data were aggregated to a monthly timescale using mean averaging.

6.2.1. Study areas

This study focuses on southeastern Australia, with the geographic extent shown in Fig. 6.1. Four representative dryland agro-ecological zones across southern New South Wales and northern Victoria were selected: Riverina rangelands (Riv), Wimmera–Mallee (Wim), Goulburn Valley (GoV), and Southwest Slopes (SwS). These regions span a gradient of land use types, including extensive grazing on native vegetation in Riv and SwS, predominantly dryland cropping in Wim, and mixed agricultural systems in GoV. Collectively, they represent key agricultural production areas and capture diverse climatic and biophysical conditions critical for drought impact assessment and early warning system development. Areas classified as irrigated land use were retained because they function predominantly as rainfed systems at the study scale and are consistent with the natural water-balance soil moisture hindcasts used in the analysis.

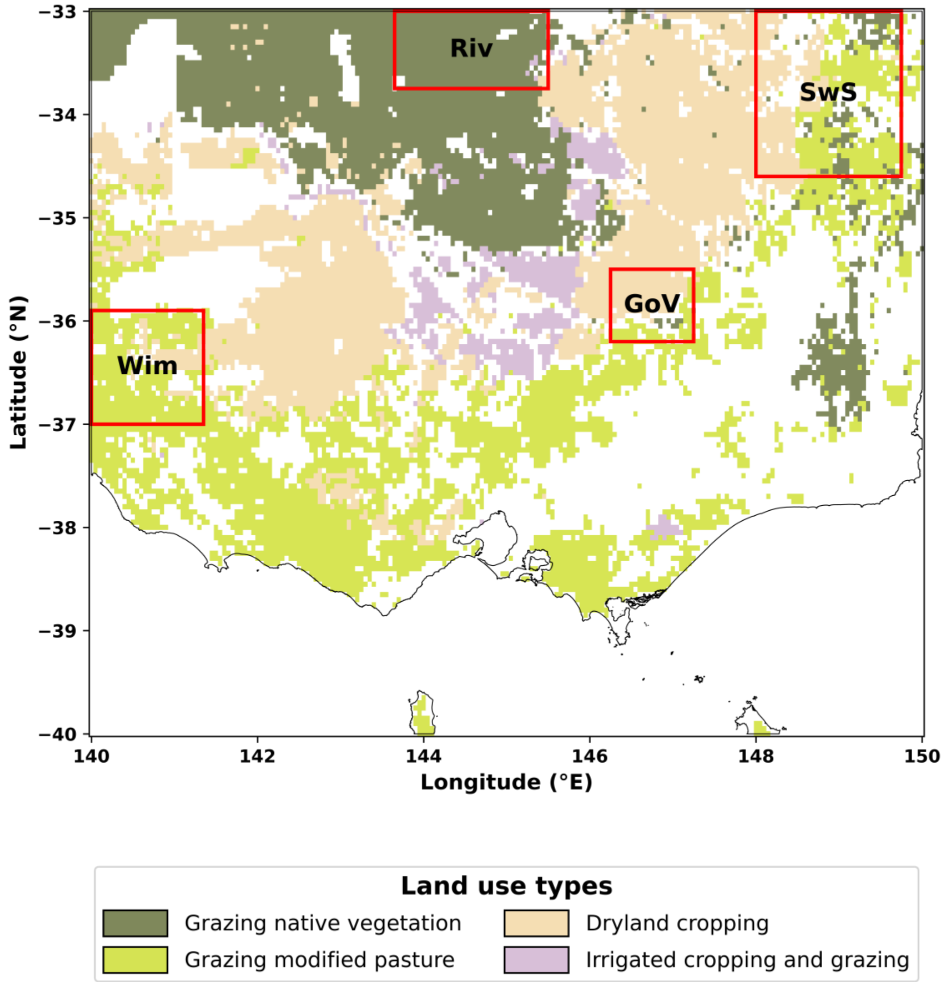


Fig. 6.1. Spatial distribution of major agricultural land use types in southeastern Australia. The red boxes delineate four key case study zones: Riverina rangelands (Riv), Wimmera–Mallee (Wim), Goulburn Valley (GoV), and Southwest Slopes (SwS). Land use types include grazing native vegetation, grazing modified pasture, dryland cropping, and irrigated cropping and grazing.

6.2.2. Ensemble soil moisture hindcast data

The daily ensemble hindcast data of root-zone soil moisture (0–1.0 m depth) used in this study were sourced from the Australian National Computational Infrastructure (NCI) Data Collection and span 1981–2017 (https://thredds.nci.org.au/thredds/catalog/iu04/australian-water-outlook/seasonal-forecasts/v1/hindcasts/catalog.html). These hindcasts were generated by forcing the AWRA-L land surface model (version 6.1) with climate forecasts from the ACCESS-S2 system, a sub-seasonal to seasonal dynamical forecasting framework developed by the Australian Bureau of Meteorology (BoM) (Hudson *et al.,* 2017). The resulting data provide root-zone soil moisture estimates at a 0.05° (~5 km) spatial resolution. AWRA-L is the BoM's operational water balance model that simulates land–atmosphere interactions to produce hydrological variables such as precipitation, soil moisture, and runoff (Frost & Shokri, 2021). The seasonal forecasting system used in this study is ACCESS-S2, BoM's current operational coupled ocean–atmosphere model, which replaced ACCESS-S1 in October 2021 (Wedd *et al.,* 2022). ACCESS-S2 produces ensemble-based seasonal climate forecasts, with each member representing an equally probable climate trajectory. Further technical documentation is available in the BoM research report (Pickett-Heaps and Vogel, 2022; http://www.bom.gov.au/research/publications/researchreports/BRR-065.pdf). AWRA-L does not explicitly simulate anthropogenic irrigation water application; therefore, the root-zone soil moisture hindcasts represent natural water-balance conditions only, including in areas classified as irrigated land use.

Each monthly hindcast cycle includes 27 ensemble members, produced using a 9-day lagged initialisation scheme consisting of three "burst" groups of nine members each. These bursts are staggered across nine consecutive start dates (e.g., 23 April to 1 May), providing lead times from 1 to 6 months. Forecasts are issued on the 1st of each month and are designed to preserve intra-seasonal signal diversity while maintaining consistency in seasonal trends (Vogel *et al.,* 2021; Pickett-Heaps and Vogel, 2022). Fig. 6.2 illustrates the configuration of this ensemble forecasting system. It shows how the hindcasts, initialised on 1st May, are used to project soil moisture conditions through the growing season.

For this study, I selected a 17-year subset (2001–2017) to align with the availability of MODIS satellite observations. Daily ensemble soil moisture hindcasts were aggregated to monthly means for each of the 27 ensemble members. Hindcasts were grouped by lead time, from 1- to 6-month lead, representing predictions issued one to six months in advance of the initialisation month. To improve normality, a square-root transformation was applied to the soil moisture data. This configuration supports the evaluation of forecast skill in predicting vegetation deterioration risks at lead times up to six months, offering early insights into vegetation conditions during the peak growing season (typically July and August; see Fig. S6.1) in southeastern Australia.

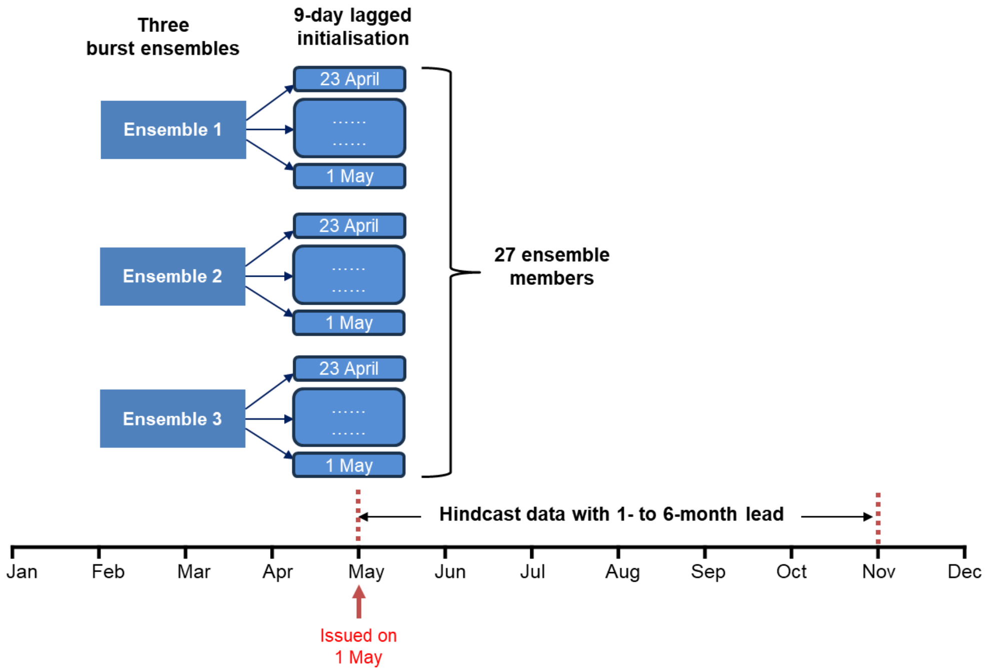


Fig. 6.2. Schematic overview of ensemble root-zone soil moisture hindcast generation. Each forecast is issued on the 1st of the month and spans lead times of 1 to 6 months (e.g., 1 May issue covers May to November). A total of 27 ensemble members are generated using a lagged initialization method, which combines three burst ensemble groups and nine staggered start dates (e.g., 23 April to 1 May).

6.2.3. Historical Soil Moisture Simulations

Historical root-zone soil moisture simulations from AWRA-L (available from 1911 to present) were used as a historical baseline to assess the forecast skill of ensemble hindcasts. These simulations, driven by observed atmospheric inputs rather than forecasted conditions, represent the "best available" estimates and serve as an idealised baseline (i.e., 0-month lead time) against which forecasts with 1- to 6-month lead times are compared. For this study, daily historical root-zone soil moisture simulations from 2001 to 2017 were sourced from the NCI Data Collection (https://thredds.nci.org.au/thredds/catalog/catalogs/iu04/australian-water-outlook/historical/historical.html) and aggregated to a monthly timescale. A square-root transformation was applied to reduce skewness and improve the normality of soil moisture values, particularly in dryland cropping and pastoral regions, thereby enhancing the robustness of the standardised anomaly analyses used in this study.

Table 6.1. Summary of the soil moisture datasets utilised in this study.

Application type	Dataset	Lead time	Simulation member number	Unit	Temporal coverage, spatial resolution
Near-real-time monitoring	Historical simulated root-zone soil moisture (Frost and Shokri, 2021)	0-month	1 member	mm/day	1911-yesterday, 0.05°
Retrospective forecast	Ensemble hindcast root-zone soil moisture (Pickett-Heaps and Vogel, 2022)	1- to 6-month	27 members	mm/day	1981-2017, 0.05°

6.2.4. MODIS vegetation greenness data

The Normalized Difference Vegetation Index (NDVI) is widely recognised as a key indicator of vegetation greenness, canopy structure, and photosynthetic activity (Huete *et al.,* 2002). In this study, the NDVI dataset was derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) nadir BRDF-adjusted reflectance product (MCD43A4 v006). The dataset was obtained from the NCI Data Collection (https://dapds00.nci.org.au/thredds/catalog/ub8/au/MODIS//mosaic/MCD43A4.006/catalog.html), and the raw 500-m resolution was resampled to 0.05° (~5 km) using bilinear interpolation for consistency with AWRA-L data. To minimise atmospheric and sensor-related noise, only the highest-quality retrievals from bands 1 and 2 were used for NDVI calculations. Temporal gaps in the NDVI time series were addressed through linear interpolation, with gap-filled data constituting only 4% of the complete dataset across the study area and analysis period. The 8-day NDVI time series were then mean aggregated into monthly intervals for the period 2001-2017.

6.3. Methodology

An overview of the experimental design and forecasting framework is presented in Fig. 6.3 and elaborated in the subsequent sections. The methodological framework in this chapter follows a two-stage design: prediction rules are firstly trained using historical root-zone soil moisture simulations and then applied to independent seasonal root-zone soil moisture hindcasts to evaluate forecasting performance by lead times. Section 6.3.1 introduces the classification of vegetation deterioration (VD) severity levels based on non-exceedance NDVI standardised anomaly thresholds. Section 6.3.2 outlines the development of the integrated VD prediction system and the configuration of optimised forecast rules derived from historical simulations. Section 6.3.3 describes the forecast verification procedure, where the derived rules are applied to SMA ensemble hindcasts to evaluate both deterministic and probabilistic skill.

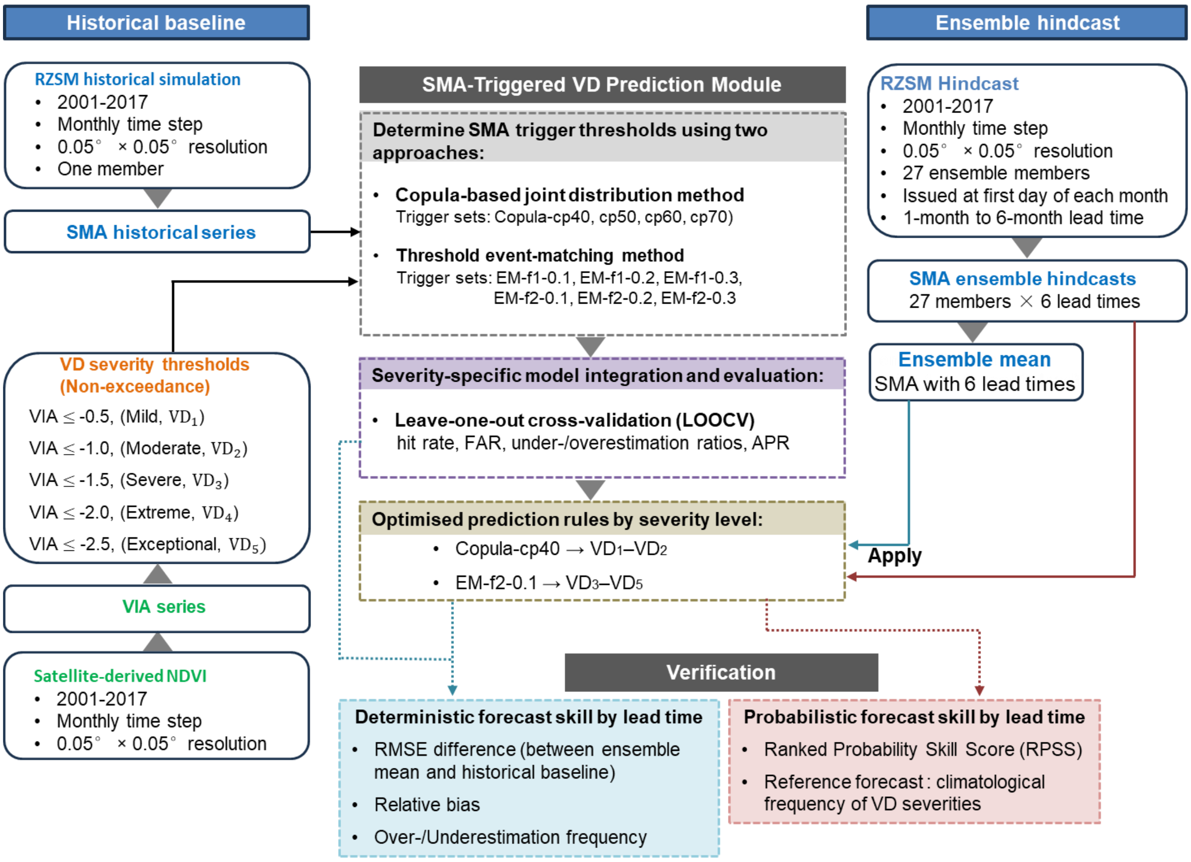


Fig. 6.3. Overview of the forecasting methodology and experimental design.

6.3.1. Classification of vegetation deterioration severity

Consistent with the methods used in Chapters 4 and 5, soil moisture deficits and vegetation stress are quantified using standardised anomalies (z-scores) of root-zone soil moisture (RZSM) and NDVI, respectively. These z-scores are computed against the 2001–2017 long-term monthly statistics following the standardised anomaly formula:

$$z_x(t, y) = \frac{x(t,y) - \mu_x(t)}{\sigma_x(t)} \tag{6.1}$$

where $x(t,y)$ is the value of the variable x on day t of year y; whereas $\mu_x(t)$ and $\sigma_x(t)$ denote the multi-year monthly mean and standard deviation of variable x on day *t,* calculated over the 2001-2017 reference period. The calculation of $x(t,y) - \mu_x(t)$ yields the anomaly for day *t* of year *y*, which is then standardised to $z_x(t,y)$. This transformation yields time-specific series of soil moisture anomalies (SMA) and vegetation index anomalies (VIA) across years and pixel-wise land use types.

In this chapter, two RZSM datasets (see Table 6.1) are utilised to derive SMA: (i) a single-member historical simulation, and (ii) a 27-member ensemble hindcast with six forecast lead times per member. Accordingly, the historical dataset yields one SMA time series, while the ensemble hindcast provides a total of 162 SMA series (27 members × 6 lead times). To ensure consistency in anomaly estimation, all SMAs, including those from the ensemble hindcasts, are calculated using the same 2001 to 2017 long-term monthly baseline derived from the historical simulation. This approach ensures that the anomalies reflect deviations from a common reference baseline, and avoids the confounding effects introduced by forecast drift or member-specific bias in the ensemble. These SMA series form the basis for determining trigger thresholds for vegetation deterioration prediction, as detailed in the following section.

Vegetation deterioration (VD) events are categorised into six severity levels to capture a gradient of stress responses. Following the classification scheme applied in Chapters 3 to 5, the severity levels are defined based on the upper-bound thresholds of VIA as follows: VD_0 (No deterioration, VIA > -0.5 SD), VD_1 (mild deterioration, VIA ≤ -0.5 SD), VD_2 (moderate deterioration, VIA ≤ -1.0 SD), VD_3 (severe deterioration, VIA ≤ -1.5 SD), VD_4 (extreme deterioration, VIA ≤ -2.0 SD) and VD_5 (exceptional deterioration, VIA ≤ -2.5 SD). Unlike conventional mutually exclusive classification systems, which define severity levels with both upper and lower bounds, this framework adopts a non-exceedance threshold approach with each severity level is independently triggered when a VIA value falls below the specified threshold. This structure is more appropriate for early warning applications, where the operational need is to determine whether a specific severity threshold has been breached, not to assign events into discrete categories. It reflects the logic of impact-based forecasting systems (WMO & GWP, 2016; Bachmair *et al.,* 2016), which rely on actionable triggers rather than exhaustive classification.

6.3.2. Integrated VD forecasting model

This section builds on the VD prediction framework established in Chapter 5 to determine SMA trigger thresholds for forecasting VD risks across different severity levels. Using the SMA timeseries derived from historical simulations as predictors, I followed the same modelling logic as in Chapter 5 to derive ten SMA trigger threshold configurations: four from the copula-based joint probabilistic models (Copula-cp40, Copula-cp50, Copula-cp60 and Copula-cp70) and six from the event-matching models (EM-f1-0.1, EM-f1-0.2, EM-f1-0.3, EM-f2-0.1, EM-f2-0.2 and EM-f2-0.3).

In Chapter 6, the total forecast horizon for VDs is defined as the combination of the soil moisture forecast lead time (L) and the vegetation response time (d), which this conceptual flaw is shown in Fig 1.2 in Chapter 1. Specifically, ensemble hindcast soil moisture provides predictions at lead times of 1–6 months ahead, while vegetation conditions respond to soil moisture anomalies with an intrinsic lag of 0–2 months, as established in Chapter 5. Accordingly, vegetation deterioration is predicted at time $T + L + d$, conditional on soil moisture forecasts issued at time T.

Although each VD severity level (VD_1 to VD_5) could theoretically select from all ten configurations, I constrained the candidate pools based on key insights from Chapter 5 to reduce computational load and enhance model interpretability: (i) copula-based models are better suited for predicting low-severity, high-frequency stress (e.g., VD_1) due to superior false alarm control; (ii) event-matching models demonstrate higher skill in identifying infrequent but severe stress events (e.g., VD_4–VD_5) via improved hit rates. Accordingly, SMA trigger configuration for VD_1 was selected exclusively from the copula-based models, for VD_4–VD_5 from the event-matching models, while VD_2–VD_3 drew from the full pool of ten SMA trigger configurations.

To construct a robust and generalisable integrated prediction model, I applied a leave-one-out cross-validation (LOOCV) procedure across the 17-year period (2001–2017). In each iteration, one year was held out as the test set, while the remaining 16 years served as the training set. Based on the training data, I reselected the optimal SMA trigger configuration for each VD severity level (VD_1–VD_5) from the constrained candidate pools. The integrated VD mapping rules was then constructed using the chosen configurations and applied to predict the VD severity category for the held-out year. This procedure was repeated for all 17 years, resulting in N=17 predicted VD maps. Model performance was evaluated by comparing these predictions against the observed VD categories for each test year, using five evaluation metrics: hit rate, false alarm ratio (FAR), underestimation ratio, overestimation ratio, and adjacent prediction ratio (APR), weighted according to severity-specific optimisation priorities (see Chapter 5, Section 5.3.3.2).

This process was repeated for all 17 years, resulting in a fully cross-validated set of optimal SMA trigger threshold configurations. Specifically, Copula-cp40 was selected for the lower-severity categories (VD_1–VD_2), while EM-f2-0.1 was assigned to the higher-severity categories (VD_3–VD_5). These integrated VD prediction rules, calibrated solely on historical SMA simulations, were then applied to the ensemble hindcast SMA series with lead times of 1 to 6 months, comprising a total of 162 series. Both deterministic forecasts (derived from ensemble means) and probabilistic forecasts (based on the full ensemble distributions) were generated and evaluated for each lead time, as detailed in the following sections 6.3.3.

6.3.3. Forecast verification: skill assessment of soil moisture hindcast

This section evaluates the forecasting skill of the ensemble-based (probabilistic) and deterministic predictions produced by the integrated framework at a monthly temporal resolution In probabilistic forecasting, a key aspect of skill assessment is whether the forecasts provide added value relative to a reference prediction, typically defined as the climatological frequency distribution of the target events (Pappenberger *et al.,* 2015). Here, the reference is constructed from the historical occurrence frequencies of each VD category (VD_0–VD_5) over the 2001–2017 period, and is referred to as the VD event climatology.

The evaluation considers both the uncertainty captured by the ensemble spread and the accuracy of deterministic predictions derived from ensemble means. Ensemble forecast skill metrics consider the ensemble uncertainty. Probabilistic forecast skill is primarily evaluated using the Ranked Probability Skill Score (RPSS), as introduced in Section 6.3.3.1. RPSS accounts for the full ensemble distribution and is appropriate for multi-category forecast verification. In parallel, deterministic forecasts derived from the ensemble means are evaluated using categorical skill metrics. The VD severity levels (VD0–VD5) are treated as discrete outcomes, and forecast performance is analysed using 6×6 contingency tables, comparing predicted categories (y_i) to observed categories (o_i). This framework allows for the computation of standard accuracy metrics such as the root-mean-square error (RMSE) and relative bias, overestimation frequency, underestimation frequency, which are defined in subsequent sections (Sections 6.3.3.2). These metrics are calculated not only at the grid-cell scale but also aggregated at the regional scale across key agricultural zones in southeastern Australia.

6.3.3.1. Probabilistic forecast skill analysis

Probabilistic forecast performance was assessed using the Ranked Probability Skill Score (RPSS), which measures the cumulative squared error between forecast categorical probabilities and observed categorical probabilities, comparing them against a reference prediction (i.e., a climatology reference prediction). RPSS is particularly well-suited for evaluating forecasts with ordinal multi-categorical severity levels where there is an imbalanced sample size of different categories (Epstein 1969). This provides a comprehensive assessment of overall forecasting performance across different VD categories.

The rank probability score (RPS) is a measure of the probabilistic forecast for a single forecast-observation pair at time *t* and is defined as (Wilks, 2011):

$$RPS_t = \frac{1}{K-1}\sum_{k=1}^{K}\left(F_{k,t} - O_{k,t}\right)^2 \ (k = 1,2,\dots,K) \tag{6.2}$$

where *K* is the total number of categories (*K*=6); *k* is the category of vegetation deterioration events (VDs); *t* = 1,…, *T* are the monthly time steps (all months or only months of a given season); the cumulative forecast probability vector at time step *t*, expressed as $F_{k.t} = \sum_{i=1}^{k} f_i$, with y_i=1 being the forecast for the event will occur in but not exceed the i-th category, and the cumulative observation vector $O_{k,t} = \sum_{i=1}^{k} o_i$ with o_i=1 if the observed event does not exceed the i-th category and o_i=0 otherwise.

The forecast skill is evaluated relative to a reference prediction, defined as the climatological occurrence frequency of VDs for each category (VD_0-VD_5) over the period 2001-2017. Let p_i be the climatological frequency of the event occur in i-th category, and let $C_k = \sum_{i=1}^{k} c_i$ be the corresponding cumulative vector. Then the climatological RPS for the reference score, RPS_{clim} at monthly time step *t* can be calculated as:

$$RPS_{clim,t} = \frac{1}{K-1}\sum_{k=1}^{K}\left(C_{k,t} - O_{k,t}\right)^2 \ (k = 1,2,\dots,K) \tag{6.3}$$

The rank probability skill score (RPSS) is defined as (Wilks, 2011):

$$RPSS = 1 - \frac{\langle RPS \rangle}{\langle RPS_{climatology} \rangle} \tag{6.4}$$

where angle brackets ⟨·⟩ denote the average of the scores over *T* forecast-observation pairs for each pixel during the annual growing season months over the period 2001-2017. The

RPSS ranges from −∞ to 1 with higher positive RPSS values indicating better forecasting skill compared to the reference (climatology) prediction. RPSS = 1 represents perfect skill and 0 indicates no skill compared to the reference.

6.3.3.2. Deterministic forecast skill analysis

To assess the deterministic forecasting skill of the ensemble mean soil moisture hindcasts in predicting vegetation deterioration (VD) events, I employed three key evaluation metrics: Root Mean Square Error (RMSE), Relative Bias (%), and Over-/Underestimation Frequency (%). These metrics provide a comprehensive assessment of forecast accuracy, systematic bias, and temporal reliability across different lead times and VD risk levels. RMSE evaluations are conducted for all risk levels (VD_0–VD_5), mild risk (VD_0–VD_2), and severe risk (VD_3–VD_5) for both the whole growing season (May–October) and seasonal scales (MJJ and ASO).

The four metrics (RMSE, relative bias, overestimation frequency, and underestimation frequency) defined below are computed for each grid cell and aggregated across southeastern Australia for different lead times (1- to 6-month lead), VD risk categories (all risk, mild risk, severe risk), and for seasons (MJJ, ASO, and full growing season).

- RMSE

RMSE is a widely used metric for assessing the accuracy of deterministic forecasts. It quantifies the average magnitude of prediction errors by calculating the square root of the mean squared differences between forecasted and observed VD categories. RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2} \quad (6.5)$$

where F_i and O_i represent the forecast and observed VD categories at time step i, respectively; and N denotes the total number of forecast-observation pairs. Lower RMSE values indicate better agreement between forecasts and observations. As there is no universally accepted threshold for a "good" RMSE, forecast skill is evaluated based on RMSE differences relative to a historical baseline. In this study, the historical baseline is defined as the RMSE computed from VD predictions generated from historical simulated SMA compared against observed VD categories. For each lead time, the difference in RMSE between hindcasts and the historical baseline (i.e., $\mathrm{RMSE}_{hindcast} - \mathrm{RMSE}_{historical}$) is used to assess the forecast performance. A negative RMSE difference indicates improvement over the historical baseline, while a positive value reflects a degradation in forecast skill.

- Relative bias (%)

Relative bias quantifies the systematic deviation of forecasts from observations, providing insight into whether the model overestimates or underestimates VD risk over the entire study period. It is calculated as:

$$\text{Relative bias} = \frac{\sum_{i=1}^{N}(F_i - O_i)}{\sum_{i=1}^{N} O_i} \times 100\% \tag{6.6}$$

A positive relative bias indicates a systematic overestimation of VD risk (e.g., a relative bias of +20% means forecasts, on average, are 20% higher than observations). Conversely, a negative relative bias signifies an underestimation of VD risk.

- Over-/Underestimation Frequency (%)

Overestimation and underestimation frequency metrics assess the temporal consistency of forecast deviations, capturing how often forecasts exceed or fall below observed VD risk levels. These metrics provide insights into whether systematic errors occur consistently across the study period or are sporadic. Overestimation frequency (%) and underestimation frequency (%) represents the proportion of time steps where the forecast VD risk category is higher and lower than observed, respectively. calculated as:

$$\text{Overestimation frequency} = \frac{\sum_{i=1}^{N} 1(F_i > O_i)}{N} \times 100\%$$

$$\text{Underestimation frequency} = \frac{\sum_{i=1}^{N} 1(F_i < O_i)}{N} \times 100\%$$

where 1(·) is an indicator function that takes a value of 1 when the condition is met and 0 otherwise. A higher overestimation frequency suggests that forecasts tend to predict vegetation deterioration events that do not occur, while a higher underestimation frequency indicates that the model frequently fails to capture observed deterioration events. For instance, if the analysis yields an overestimation frequency of 20% and an underestimation frequency of 40%, this indicates that the forecast system produces values exceeding observations in 20% of temporal instances and falls below observations in 40% of temporal instances.

6.3.3.3. Aggregation of forecast metrics at regional scales

In addition to grid-based verification, forecast skill metrics were aggregated at the regional scale to evaluate performance across the four case study zones outlined in Fig. 6.1: Riverina rangelands (Riv), Wimmera–Mallee (Wim), Goulburn Valley (GoV), and Southwest Slopes (SwS). These regions represent distinct agro-ecological contexts and land use systems across southeastern Australia. For each case study zone and lead time, spatial averages of RMSE differences and RPSS values were computed to assess the skill of vegetation deterioration (VD) forecasts. This regional-scale aggregation enables a more interpretable comparison of forecast performance in relation to land management zones.

6.4. Results

6.4.1. Performance of the probabilistic VD forecast

6.4.1.1. Ranked probability skill score

I assessed the accuracy and reliability of ensemble soil moisture forecasts in predicting ordinal multi-categorical VDs through the RPSS, calculated across agricultural land grid cells in southeastern Australia for the period 2001-2017. The RPSS measures relative improvement of the forecast performance against a climatology forecast, where positive values indicate superior performance compared to climatology ensemble forecasts. The RPSS values near zero or negative indicate no improvement or diminished performance, respectively.

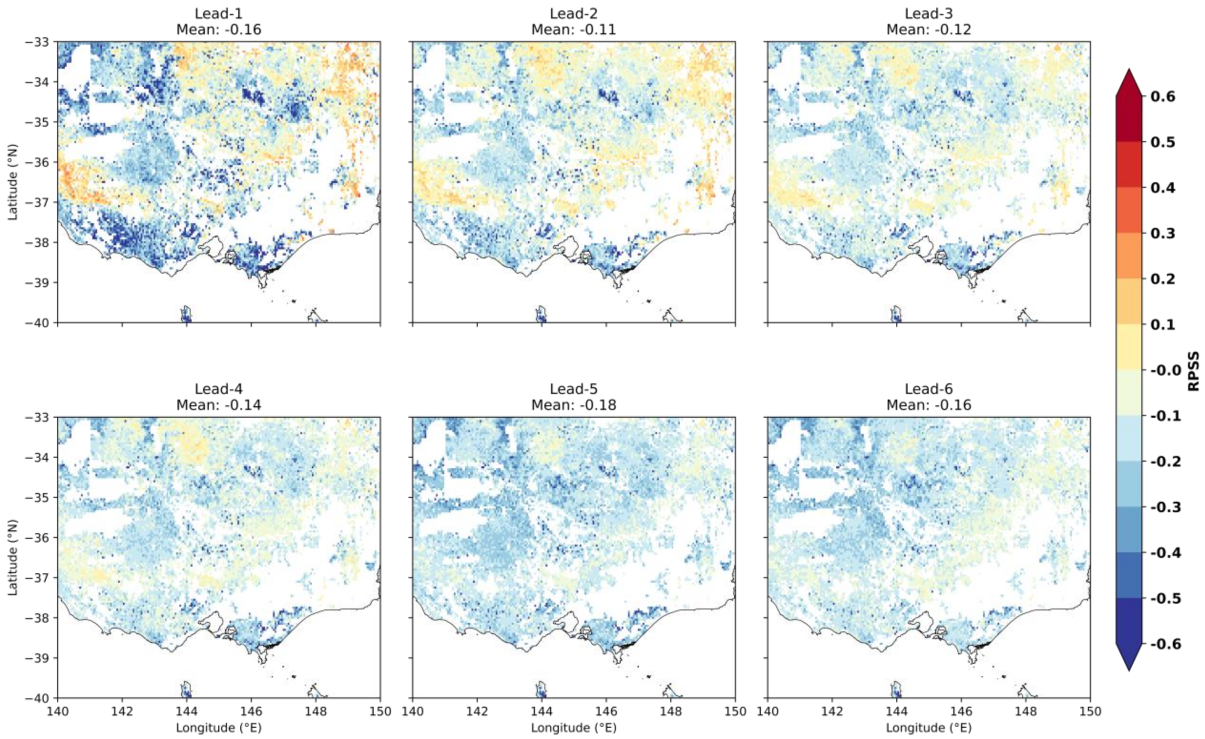


Fig. 6.4. Spatial distribution of Ranked Probability Skill Score (RPSS) values for vegetation deterioration (VD) forecasts at 1- to 6-month lead times across southeastern Australia, for the seasonal periods of the growing season (May to October) during 2001–2017. Warmer colours (redder shades) indicate higher forecast skill.

Fig. 6.4 illustrates the spatial distribution of RPSS values throughout the growing season (May to October) at varying lead times (1- to 6-month lead). A detailed seasonal breakdown of RPSS, distinguishing between early (May-June-July, MJJ) and late (August-September-October, ASO) growing seasons is provided in Supplementary Fig S6.2. Comprehensive RPSS results at seasonal and key agricultural regional scales are presented in Appendix Tables S.6.6. Over the whole growing season, the RPSS values show a positive skill in specific regions of southeastern Australia, with the area proportions of positive RPSS for 1-month to 6-month lead being 21.2%, 21.5%, 9.4%, 4.1%, 0.4%, and 0.7%, respectively. These results demonstrate that the predictive capability of ensemble soil moisture hindcasts exhibits a clear degradation at extended lead times.

The temporal analysis of RPSS reveals pronounced seasonal variations in forecast skill (see Fig. S6.2). During the late growing season (ASO), the framework demonstrates superior predictive capability over that for the early growing season (MJJ) and maintains positive RPSS values up to 2-month lead across all regions, with some areas exhibiting skill extending to 3-month lead (Fig. S6.2). The ASO period exhibited both higher and widespread positive RPSS values than that of MJJ, with the corresponding area proportions for the 1-month to 6-month lead being 41.3%, 48.4%, 35.0%, 24.3%, 12.8%, and 4.9%, respectively. Conversely, the MJJ season exhibits limited regions with positive RPSS values; the spatial extent of skilful forecasts for the 1-month to 6-month lead are 11.6%, 8.0%, 1.1%, 0.5%, 0.0%, and 15.1%, respectively, highlighting that skillful forecasts are restricted to a few areas in the early part of the growing season. This pattern suggests fundamental challenges in predicting vegetation deterioration events during initial growth phases.

Interestingly, the framework shows the highest mean RPSS (see panel headings on Fig. 6.4) at the 2-month lead across southeastern Australia. While the 1-month lead exhibits higher positive RPSS values in some regions, more negative RPSS values are observed in coastal grazing modified pasture areas that result a lower overall mean RPSS at the 1-month lead across the study region. This spatial pattern is more pronounced during the ASO season, as shown in Fig S6.2g. In addition, the ensemble VDs forecasting framework exhibits significant spatial variations across southeastern Australia. In general, the forecast skill tends to be higher in parts of the northeastern, western, and central regions, particularly in areas dominated by grazing native vegetation and grazing modified pasture land use types maintaining positive skill metrics across 1- to 3-month lead. From a seasonal perspective, while forecast skill is initially constrained earlier in the cool season, a systematic enhancement emerges as the season progresses into the ASO period, with predictive capability extending to a 4-month soil moisture lead in some regions.

6.4.1.2. Skill across total forecast horizons: combined vegetation response and extended lead times

The VD forecasting framework accounts for both the intrinsic vegetation response time to soil moisture anomalies (d) and the lead time of ensemble soil moisture hindcasts (L). In this framework, vegetation deterioration is predicted at time $T + L + d$, such that the total forecast horizon reflects the combined effect of soil moisture forecast lead time and vegetation response lag. Fig. 6.5 summarises forecast skill as a function of soil moisture hindcast lead time, conditional on different assumed vegetation response times.

The boxplots in Fig. 6.5a show substantial spatial variability in RPSS at a 1-month soil moisture lead, with some pixels exhibiting positive skill and others showing strongly negative values. This spatial heterogeneity is particularly evident for pixels characterised by a 2-month vegetation response time, where RPSS values at a 1-month soil moisture lead can be lower than those at longer lead times. This pattern reflects the interaction between intrinsic system delays and forecast uncertainty when short-lead soil moisture forecasts are used to predict vegetation conditions at longer effective horizons.

The spatial distribution of vegetation response times (Fig. 6.5b) indicates that most agricultural land exhibits a 1-month response to soil moisture variability (79.8%), while smaller proportions show immediate (0-month, 13.4%) or delayed (2-month, 6.7%) responses. Across increasing soil moisture hindcast lead times, median RPSS values change only modestly, whereas the overall spread of RPSS values narrows. In particular, the most positive RPSS values decline with lead time, while strongly negative values become less extreme, resulting in reduced variability in forecast skill distributions.

Negative RPSS values primarily arise from the aggregated probabilistic performance across all VD severity categories (VD_0–VD_5), reflecting underprediction at low VD risk levels and overprediction at high VD risk levels. Despite predominantly negative RPSS values across much of the study area, the spatial and temporal structure of forecast skill highlights how forecast uncertainty and intrinsic vegetation response times jointly shape predictability across extended forecast horizons.

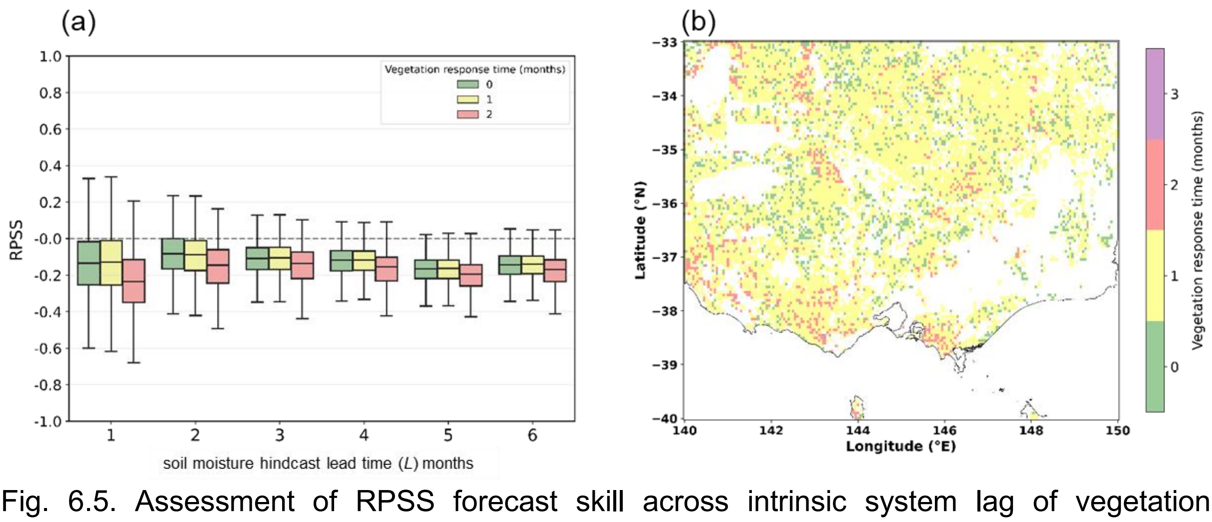


Fig. 6.5. Assessment of RPSS forecast skill across intrinsic system lag of vegetation response to soil moisture (0-2 months, *d*) and extended (1- to 6-month) lead times (*L*) from ensemble soil moisture hindcasts. (a) Distribution of Ranked Probability Skill Score (RPSS) across vegetation response times (0-2 months) for soil moisture hindcasts at 1- to 6-month lead times, presented as box-and-whisker plots. (b) Spatial distribution of vegetation response times implemented in the prediction framework across southeastern Australia. The analysis reveals that 13.4%, 79.8%, and 6.7% of agricultural land exhibits 0-month, 1-month, and 2-month vegetation response times to soil moisture variations, respectively.

6.4.2. Performance of the deterministic VD forecast

The performance of deterministic VD forecasts from ensemble mean SMA was evaluated using three groups of metrics: RMSE difference, relative bias (%), and over-/underestimation frequency (%). These metrics were assessed across three VD risk intervals: all risk levels (VD_0–VD_5), mild risk (VD_0–VD_2), and severe risk (VD_3–VD_5). Evaluations were conducted for both the full growing season (May–October) and two sub-seasonal (MJJ and ASO).

As shown in Fig. 6.6 for the full May–October period, RMSE differences relative to the historical baseline vary across lead times and regions. At the 1-month lead, negative RMSE differences are observed across many regions, indicating locally improved performance relative to the baseline. At the 2-month lead, RMSE differences are close to zero, suggesting performance broadly comparable to the historical baseline. Beyond the 3-month lead, RMSE differences become predominantly positive, with regional mean values increasing to approximately 0.08–0.09, indicating degraded forecast performance relative to the baseline. The early cool-season period (MJJ; Fig. S6.3a) exhibits more spatially extensive negative RMSE differences than the full-season average, indicating relatively improved forecast performance during this period. In contrast, the late-season period (ASO; Fig. S6.3b) shows greater spatial heterogeneity and more pronounced degradation in forecast skill at longer lead times. A notable exception occurs in southern coastal regions dominated by grazing on modified pasture, where forecasts at 3- to 6-month lead exhibit consistently lower RMSE than the historical baseline.

As shown in Tables S6.1–S6.4, the relative bias analysis reveals that the deterministic forecast framework systematically underestimates VD risk, with negative bias intensifying as lead time increases.

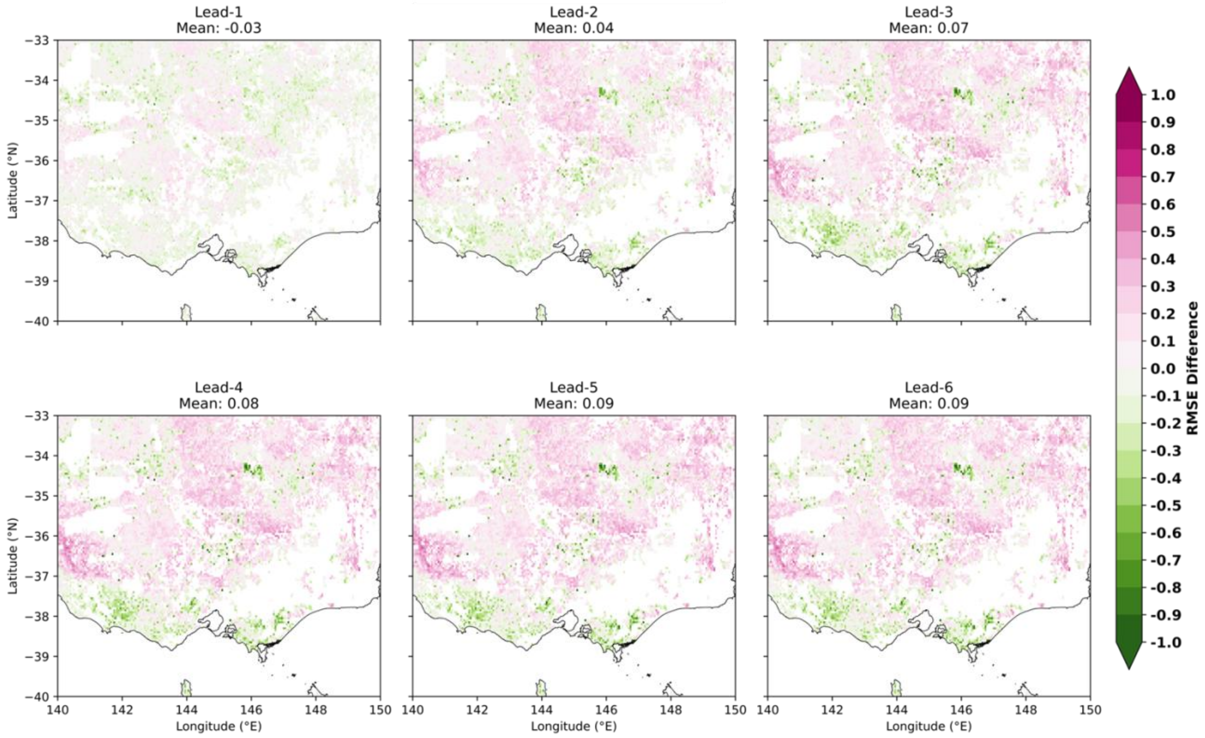


Fig. 6.6. Spatial distribution of deterministic forecast skill for vegetation deterioration (VD) events across southeastern Australia, shown for the full growing season (May–October). Forecast skill is represented by the difference in root mean square error (RMSE) between hindcasts and the historical baseline (i.e., $\mathrm{RMSE}_{hindcast} - \mathrm{RMSE}_{historical}$) for 1- to 6-month lead times during 2001–2017. The mean RMSE difference, spatially averaged across the study region, is reported above each panel. Negative values (green) indicate improved forecast skill relative to the historical baseline, while positive values (pink) indicate degradation.

From a risk interval perspective, while the historical baseline shows minimal overall bias, it exhibits clear deviations within specific risk intervals, with a -10% bias for mild risk and a +32% bias for severe risk. This asymmetry is particularly pronounced during the MJJ season, where the bias for severe risk categories reaches approximately +70%. Across all lead times and growing season segments, deterministic forecasts show consistent bias patterns within mild and severe risk categories. Forecast accuracy deteriorates progressively, with average bias shifting from –20% at 1-month lead to –44% at 6-month lead. In terms of bias occurrence frequency, the historical baseline tends to underestimate in ~30% of cases and overestimate in ~24%, whereas the deterministic forecasts exhibit a systematic shift toward underestimation. The underestimation frequency increases from 36.2% at 1-month lead to 43.0% at 6-month lead, while overestimation decreases from 12.9% to 1.5%. This underestimation tendency becomes particularly prominent for severe risk during the late growing season (ASO), where underestimation frequencies exceed 60%. Conversely, overestimation drops sharply beyond 1-month lead, especially in the mild and all-risk categories.

Overall, these analytical outcomes demonstrate that deterministic forecasts maintain robust performance up to 2-month lead times, particularly within central and northern geographical domains. However, forecast accuracy exhibits systematic degradation beyond 3-month leads, characterised by persistent underestimation bias that manifests most prominently in severe vegetation deterioration events? during the early growing season.

6.4.3. Forecast performance over regions important for agriculture

This section examines forecast performance within four pre-defined agricultural analysis regions in southeastern Australia, selected to represent distinct wheat–sheep production environments (Riverina rangelands, Wimmera–Mallee, Goulburn Valley, and Southwest Slopes; see Fig. 6.1). These regions are used as case-study subdomains to explore how forecast skill varies across contrasting agricultural systems, rather than to represent average conditions across the broader agricultural landscape.

The RMSE difference and RPSS averaged over four wheat-sheep producing regions are shown in Fig. 6.7 (see Tables S6.5 and S6.6 in the Supplementary Information for results at the seasonal scale). Across all regions, RMSE difference increases with forecast lead time, indicating a progressive decline in deterministic forecast accuracy relative to the historical baseline. For the full May-October period (Fig. 6.7a), RMSE differences are close to zero at the 1-month lead and become increasingly positive beyond 2-month lead, stabilising at approximately 0.2-0.3 from 3-month lead onward. This indicates that sub-seasonal deterministic forecasts (up to 2-month lead) perform broadly comparable to the historical baseline but exhibit increasing error at longer leads. A similar pattern is observed during MJJ period (Fig. 6.7b), with generally lower RMSE differences than the full-season average. In contrast, RMSE differences during ASO (Fig. 6.7c) increase more rapidly with lead time, exceeding 0.3 at 6-month lead in regions such as Wimmera–Mallee and Goulburn Valley.

Relative to southeastern Australia–wide averages (Fig. 6.6), forecast performance within the selected analysis regions shows greater variability across both lead time and season (Fig. 6.7d–f). This reflects the fact that these regions represent subsets of the broader agricultural landscape and include areas where forecast skill is locally higher or lower than the regional mean.

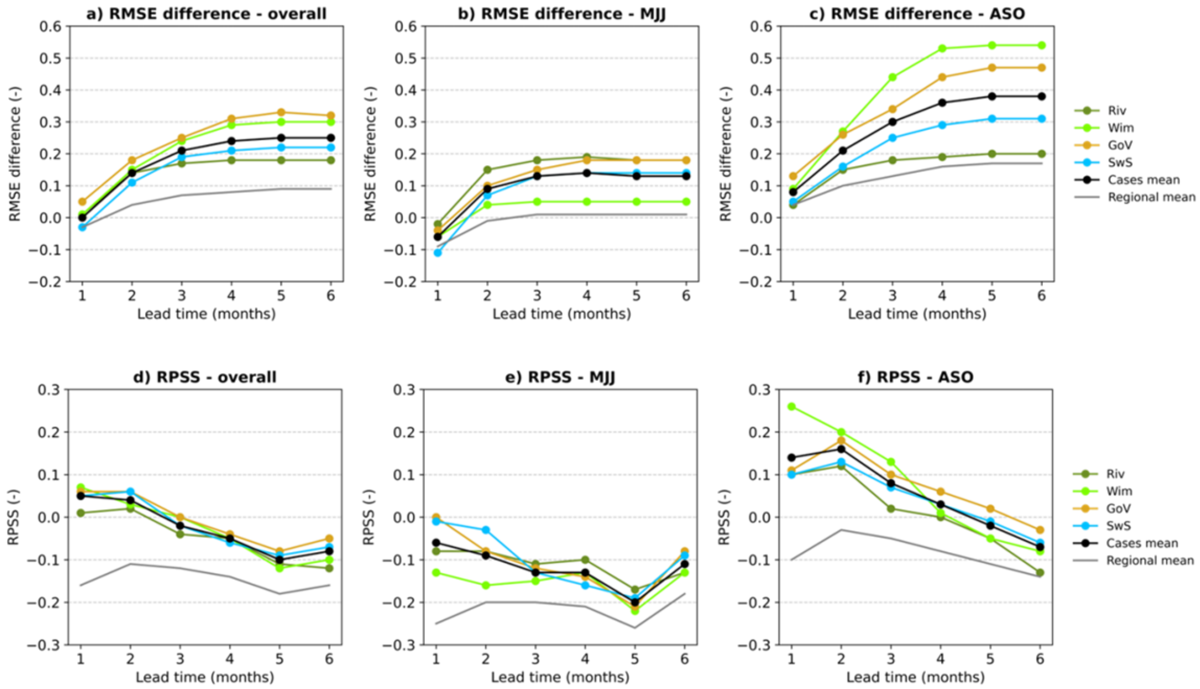


Fig. 6.7. RMSE difference (a-c) and RPSS (d–f) averaged over 4 key agricultural zones in southeastern Australia – for vegetation deterioration events (VDs) in all growing season months (a and d), MJJ (b and e), and ASO (c and f) (see Fig. 6.1 for a map of the key agricultural zones across the study region). The black lines and markers represent the mean across these 4 key agricultural zones. Grey lines represent the southeastern Australia-wide averages using gridded data for each skill metric. Riv = Riverina rangelands, Wim = Wimmera-Mallee, GoV = Goulburn Valley, SwS = Southwest Slopes.

For the full growing season (Fig. 6.7d), the mean RPSS across the selected regions is close to zero at the 1-month lead and declines with increasing lead time, becoming negative beyond 3–4 months. This indicates limited probabilistic forecast skill relative to climatology at longer leads. Seasonal differences are evident: RPSS values during MJJ are negative across all selected regions at all lead times (Fig. 6.7e), whereas during ASO, RPSS values are closer to zero or weakly positive in some regions, particularly Wimmera–Mallee and Goulburn Valley, at lead times up to 4–5 months (Fig. 6.7f).

Overall, these results indicate that forecast performance within selected agricultural regions can differ from the southeastern Australia–wide average, but positive skill is spatially and seasonally limited. The results highlight where and when soil moisture anomaly–based forecasts may provide comparatively better information within agricultural systems, while also underscoring the prevalence of negative skill across much of the broader landscape.

6.4.4. Case study of VDs in 2002

I examine the vegetation deterioration events (VDs) in August 2002 over southeastern Australia as a case study to illustrate the behaviour of the proposed forecasting framework under a well-documented drought event. The temporal evolution of VIA throughout 2002 (Fig. 6.8c) shows a progressive regional decline in vegetation condition beginning in May, with peak deterioration occurring during the late growing season in August. Severe VDs were predominantly observed across the central and northern regions (Fig. 6.8a). The VDs predicted via historical soil moisture simulation show reasonable agreement with observations (Fig. 6.8b), though notable underestimation occurred in the western and southwestern regions.

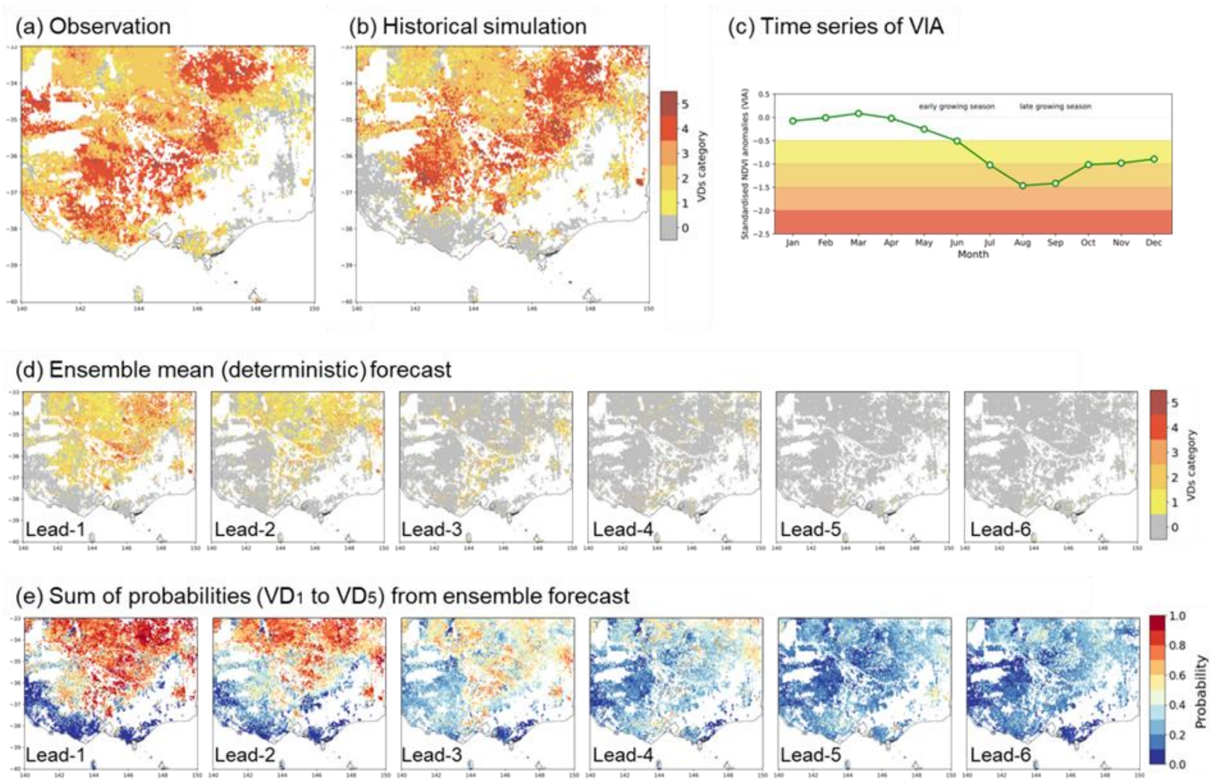


Fig. 6.8. Observation and forecasts of vegetation deterioration events (VDs) across southeastern Australia for August 2002. (a) Observed VDs categories; (b) VDs categories predicted using historical soil moisture simulation with zero lead time; (c) Time series of regionally averaged Vegetation Index Anomaly (VIA) from January to December in 2002, displaying early and late growing seasons; (d) Deterministic forecasts of VD categories for August 2002, derived from ensemble mean soil moisture at lead times ranging from 1- to 6-month; (e) Probabilistic forecast VD occurrence for August 2002, based on ensemble soil moisture forecasts issued at lead times ranging from 1- to 6-month.

Deterministic forecasts based on ensemble mean soil moisture illustrate how forecast skill varies with lead time for this event (Fig. 6.8d). At short subseasonal lead times (1–2 months), the deterministic forecasts broadly capture the spatial extent of observed VD-affected areas, particularly in the northern part of the study region. However, the severity of vegetation deterioration is not well reproduced, and forecast skill declines substantially at longer lead times (3–6 months), with limited detection of VD signals.

Probabilistic forecasts derived from ensemble soil moisture outputs represent the forecasted probability of VD occurrence aggregated across VD_1–VD_5 categories (Fig. 6.8e). At lead times up to 3 months, higher forecast probabilities are concentrated in regions that experienced observed VD impacts, while forecasted probabilities decrease systematically with increasing lead time, becoming markedly lower at 5–6 months. Comparison of accumulated probabilities for mild (VD_1–VD_2) and severe (VD_3–VD_5) VD categories (Fig. S6.4) indicates that, at short lead times, the ensemble forecasts assign relatively higher probabilities to severe VD outcomes in northern regions, highlighting the framework's ability to differentiate between risk intervals for this event.

Overall, this case study demonstrates that the proposed framework can reproduce key spatial features of a major drought event at short lead times and illustrates the progressive degradation of forecast information with increasing lead time. Additional drought-year case studies presented in the Supplementary Information (e.g. October 2006; Fig. S6.5) show comparable forecast behaviour, reinforcing the broader applicability of the framework for characterising vegetation drought risk evolution rather than providing event-specific skill validation.

6.5. Discussion

6.5.1. Framework uncertainties and limitations

This framework explores the potential of using soil moisture forecasts to anticipate vegetation deterioration events at seasonal timescales across southeastern Australia by extending the inherent lagged relationship between soil moisture and vegetation response. While the results demonstrate that root-zone soil moisture provides a physically meaningful basis for linking hydroclimatic conditions to vegetation dynamics, several methodological and data-related limitations warrant careful consideration.

First, the framework relies on root-zone soil moisture as the primary predictor of vegetation deterioration. Although soil moisture represents a key integrative variable governing vegetation water availability, vegetation anomalies may also arise from factors not explicitly represented in the framework, including land management practices, pest outbreaks, and wildfire disturbances. These processes can influence vegetation condition independently of soil moisture deficits and may therefore contribute to unexplained variability in the predicted vegetation responses.

Second, the analysis is constrained by the 17-year study period (2001–2017), which limits the sample size available for evaluating forecast skill, particularly for rare and high-severity vegetation deterioration events. This limitation primarily reflects data availability rather than a structural weakness of the proposed framework and is common to many seasonal drought forecasting studies. Nevertheless, small sample sizes may reduce the stability of skill metrics for extreme categories, and caution is therefore required when interpreting performance statistics for severe vegetation deterioration.

Third, the framework assumes independence of drought impacts between years and does not explicitly account for legacy effects associated with multi-year climate variability. This simplification may lead to underestimation of threshold events and misidentification of drought impact onset dates, particularly in scenarios where soil moisture deficits persist from pre-growing seasons. Previous studies have shown that large-scale climate modes such as ENSO, as well as their modulation by the Interdecadal Pacific Oscillation, can influence soil moisture persistence and vegetation responses across seasons and years in Australia (Power *et al.,* 1999; Donohue *et al.,* 2009)

Finally, the framework establishes soil moisture–vegetation relationships using growing-season data (May–October), while early-season forecasts necessarily rely on soil moisture conditions from preceding months. This temporal mismatch does not constitute a fundamental limitation of the framework, but it may influence forecast performance if soil moisture forecasts exhibit seasonally varying skill or if transitional processes at the onset of the growing season are not fully captured. A more explicit assessment of soil moisture forecast skill across seasons would help to disentangle the relative contributions of predictor uncertainty and vegetation response dynamics, and represents an important direction for future work.

6.5.2. Seasonal and spatial forecast skill variation

The framework exhibits distinct seasonal patterns in forecast skill, with higher performance generally observed during the late growing season (August–September–October, ASO) compared to the early growing season (May–June–July, MJJ). This seasonal contrast may reflect differences in the dominant hydroclimatic drivers across the growing season. Cool-season rainfall, which is often associated with soil moisture anomalies from preceding autumn conditions (March-May), typically intensifies across southeastern Australia during early winter and reaches maximum influence in late winter (August) before declining in spring (October) (Wedd *et al.,* 2022). Lower forecast skill during MJJ may arise from several interacting processes. One possible explanation is the influence of agricultural management activities, such as sowing and early crop establishment, which can modify vegetation dynamics independently of soil moisture availability during early-season transition periods (Chenu *et al.,* 2013; Feng *et al.,* 2020). In addition, vegetation dormancy and rapid phenological transitions during this period may weaken the strength of the soil moisture–vegetation coupling inferred from growing-season data.

Forecast skill also exhibits pronounced spatial variability, particularly at short soil moisture lead times. At the 1-month lead, spatial heterogeneity contributes to lower area-averaged skill scores than those observed at the 2-month lead (Fig. 6.4). This contrast becomes more evident during the ASO period (Fig. S6.2). Interestingly, even within the selected analysis subdomains located within four key agricultural zones, where positive RPSS values are concentrated, two regions exhibit lower average RPSS at 1-month lead compared to 2-month lead forecasts (Fig. 6.4d). This pattern contrasts with theoretical expectations and empirical evidence from various hydro-meteorological forecasting applications (e.g., precipitation, soil moisture, runoff), where shorter lead times typically demonstrate superior forecasting skill (van Dijk, 2013; Vogel *et al.,* 2021). However, given the limited spatial extent of the analysis subdomains, it remains unclear whether this apparent reduction in skill represents a robust signal or reflects sampling uncertainty.

The spatially heterogeneous skill patterns (Fig 6.8b) suggest that forecast performance is strongly conditioned by local environmental characteristics. For example, in coastal regions south of the Grampians in Victoria, which include mixed grazing and dairy production systems, the framework exhibits lower probabilistic skill relative to climatology, despite comparatively smaller RMSE values. This contrast may reflect complex land–atmosphere interactions in coastal environments, where maritime influences from the Southern Ocean and Tasman Sea, combined with intensive land use, introduce substantial variability in vegetation indices that is not fully captured by soil moisture anomalies alone (Dey *et al.,* 2019).

In some locations, reduced forecast skill may also reflect a mismatch between the inferred vegetation response time and the effective soil moisture–vegetation dynamics operating locally. Although response times are estimated empirically from historical data, their spatial generalisation may be limited in environments where vegetation responses are strongly influenced by non-hydrological controls. This interpretation is consistent with the relatively weak soil moisture–vegetation correlations identified in coastal grazing systems (Chapter 3, Fig. 3.2).

Overall, the pronounced spatial variability in forecast skill indicates that limitations of the framework are not solely attributable to insufficient regional tuning, but may instead reflect more fundamental constraints in representing vegetation responses using soil moisture alone across diverse environments. These findings highlight the need to better characterise the conditions under which soil moisture provides a dominant control on vegetation dynamics, as well as the circumstances in which additional processes must be considered.

6.5.3. Total forecast horizons and added value of impact-based drought forecasting

The total forecast horizon of the proposed framework comprises two key components: (i) the vegetation response time (typically 0-2 months), which represents the lagged response of vegetation to soil moisture anomalies, and (ii) the lead time of soil moisture forecasts available from the forecasting system, which extends beyond the contemporaneous soil moisture–vegetation relationship. These two components, vegetation response time (*d*) and forecast lead time (*L*), are conceptually illustrated in Fig. 1.2b. Together, these components support a subseasonal-to-seasonal early warning capability with an effective forecast horizon of approximately 2–4 months for anticipating vegetation deterioration risk, although useful skill is spatially limited and confined to specific regions.

Model training and verification were conducted using historical soil moisture simulations (see Chapter 5, Section 5.3.4; Chapter 6, Section 6.3.2), with a deliberate emphasis on favouring overestimation of drought impacts. This conservative design choice is motivated by the need to minimise missed warnings in high-stakes agricultural settings, where failure to detect vegetation stress may result in substantial economic losses (Stone & Meinke, 2006). Nevertheless, analysis reveals that deterministic forecasts derived from ensemble mean soil moisture hindcasts tend to significantly underestimate vegetation deterioration risk (Fig. 6.8d).

Probabilistic forecasts, when assessed using proper scoring rules, offer a decision-relevant framework for impact-based early warning, especially where forecast performance is heterogeneous across space (Arnal *et al.,* 2018). In this context, applying stakeholder-specific probability thresholds may improve the practical utility of forecasts. For instance, as illustrated in Fig. 6.8e, stakeholders may consider initiating drought preparedness measures when the forecast probability of moderate vegetation deterioration (VD_2) exceeds 70%, and begin raising awareness when the forecast probability of extreme deterioration (VD_4) risk exceeds 30%.

While challenges remain in precisely forecasting the timing and intensity of drought impacts, this impact-based forecasting framework provides a structured way to explore how soil moisture information can be translated into probabilistic indications of vegetation deterioration risk. Rather than directly enabling early intervention, the framework may help to contextualise drought risk signals and support exploratory risk awareness, particularly in drought-vulnerable regions (Murphy, 1993; Arnal *et al.,* 2018). At the local level, the forecasts can assist agricultural stakeholders in adjusting grazing practices, assessing crop viability, and managing resources. This includes enabling pastoralists to prepare targeted drought response strategies such as emergency water provision, supplementary feeding, livestock adjustment, and relocation planning (FAO, 2019). At the regional level, the framework may support integrated risk assessment by aggregating local warnings to quantify spatial patterns of drought severity. Such assessments could assist state and federal authorities in evidence-based drought response planning and may provide opportunities to align with economic indicators for broader impact forecasting.

6.5.4. Implications and future research directions

The evaluation of the inherent performance of the historical baseline for vegetation deterioration (VD) prediction highlights important constraints on the predictive system and indicates directions for future refinement. While the integration of copula-based and event-matching models provides a useful foundation for multi-category risk prediction, forecast performance remains uneven across regions and severity levels. These limitations may reflect not only aspects of the framework structure, but also the relatively low information content available in the predictor variables, particularly when soil moisture forecasts exhibit limited skill or weak coupling with vegetation response. Future refinements could therefore focus on both methodological development and improved utilisation of available information. These may include hybrid modelling strategies, such as dynamic threshold adjustment that accounts for evolving climate conditions and spatial heterogeneity (De Kauwe *et al.,* 2020). A key enhancement would involve spatially explicit weighting schemes to optimise the contribution of each model type based on region-specific skill metrics.

In addition, incorporating additional predictors such as antecedent soil moisture anomalies and initial vegetation stress conditions may help improve forecast accuracy, particularly in regions where VD events in early growing season or moderate VD events currently tend to be underestimated. These enhancements have the potential to strengthen the framework's practical relevance and responsiveness to varying drought risk contexts.

The current geographical scope of the analysis focuses southeastern Australia (primarily across Victoria, the Australian Capital Territory, southern New South Wales and a small portion of eastern South Australia). However, the analysis has yet to explore several major agricultural production regions across the continent, including the intensive farming zones in Queensland and the coastal agricultural districts of South Australia (SA) and Western Australia (WA). Future research endeavours could aim to expand this forecasting framework to encompass the entire continental agricultural landscape, extending from the winter-dominant rainfall regions in the southwest to the summer-dominant northern agricultural zones (Bruno *et al.,* 2018). This continental-scale implementation would provide comprehensive coverage across the diverse farming systems within Australia, each characterised by distinct climatological patterns, soil conditions, and agricultural practices (Hochman *et al.,* 2017). Such expansion would enhance the utility of the framework for national agricultural planning and drought risk management strategies.

The transition from hindcasts to operational forecasting introduces significant methodological adaptations and opportunities for advancement. The current analysis employed hindcast data at monthly timescales (2001-2017), whereas the operational Australian Water Outlook (AWO) system implements a 99-member time-lagged ensemble for real-time forecasts (available from October 2021, https://awo.bom.gov.au/about/overview/forecasts). This represents a substantial expansion beyond the 27-member hindcast ensemble members examined in this study. The real-time operational seasonal forecasts within AWO deliver soil moisture estimation through a set of ensemble percentiles (5th, 25th, 50th, 75th and 95th) at lead times of 1- to 3-months. Future research includes leveraging these soil moisture percentile forecasts into the established VDs prediction framework.

From a socioeconomic perspective, future research directions would evaluate the implications of impact-based forecasting by quantifying the cost-benefit trade-offs for drought preparedness interventions. A critical advancement would involve linking VDs forecast outputs to economic valuation frameworks to examine drought-associated costs through relative value analysis, particularly focusing on cost-loss ratios for optimising early action trigger thresholds (Busker *et al.,* 2023; Richardson, 2000; Wilks, 2001). This integration would support the development of tailored impact forecast products that address the diverse manifestations of soil moisture drought-induced vegetation stress, including reduced crop yields, ecological degradation and enhanced fire risk (Crausbay *et al.,* 2017).

6.6. Conclusions

This chapter evaluated the predictive capability of seasonal ensemble soil moisture hindcasts for detecting drought-induced VDs across agricultural regions in southeastern Australia from 2001 to 2017. It extended and implemented an integrated prediction framework that synthesises copula-based and event-matching methods to identify optimal SMA trigger thresholds for categorising multi-level VD risks. The framework first established SMA threshold configurations using historical simulations of root-zone soil moisture as historical baseline, and then validated their transferability using ensemble hindcasts at seasonal lead times ranging from 1 to 6 months.

The findings indicate that ensemble soil moisture hindcasts offer context-dependent and spatially variable information for forecasting drought-induced VDs, with performance dependent on land use type, season, and forecast lead time. Key insights from this study are as follows:

1. Total anticipate horizon of VD forecasts spanned about two to four months – this window comprised two components: the intrinsic vegetation response lag to soil moisture deficits (typically 0–2 months) and the extended forecast lead time from seasonal soil moisture hindcasts (1- to 3-month lead). Combined, these factors enabled early warning capability up to four months in many regions of southeast Australia. Forecast accuracy declined substantially beyond this window, especially for severe deterioration categories.

2. Probabilistic and deterministic forecasts provided complementary information for VD risk assessment – these two approaches were not directly compared, as they serve complementary roles in early warning. While RPSS-based evaluation showed that probabilistic forecasts exhibited limited category-level skill beyond the 1-month lead in most regions, they demonstrated an ability, as illustrated in the 2002 case study, to distinguish between two broader risk intervals: mild (VD_1–VD_2) and severe (VD_3–VD_5). Case studies revealed that where the forecast probability of a given risk interval exceeded 60% at the 2-month lead, the corresponding VD outcome was often confirmed by observations. Deterministic forecasts, although more straightforward and easier to interpret, tended to substantially underestimate the occurrence of severe risk interval (VD_3–VD_5) beyond the 2-month lead. Taken together, these two forecasting modes offer complementary strengths for operational decision-making under different levels of risk tolerance.

3. Forecast skill exhibited clear spatial and seasonal variation – higher skill was observed in inland pastures and dryland cropping regions, while coastal pasture regions showed relatively poor performance. Seasonally, forecasts were generally more accurate during the late growing season (August–October) than the early season (May–July). These spatial and temporal patterns highlight the need for region- and season-aware interpretation of forecast outputs in operational applications, rather than uniform use of forecast information.

4. The framework added value for impact-based drought early warning at multiple spatial scales – by translating soil moisture anomalies into multi-categorical VD warnings and aggregating forecasts across space, the system provided timely and scalable information. At the local scale, the forecasts could support decision-making by farmers and pastoralists around grazing, irrigation, and resource planning. At the regional level, they could inform broader drought risk assessments and contribute to state or federal preparedness strategies.

In summary, the findings in this chapter show that ensemble soil moisture hindcasts can provide temporally extended but spatially heterogeneous information relevant to anticipating vegetation deterioration risk several months in advance. The primary contribution of this work lies in demonstrating how seasonal soil moisture forecasts can be structured and translated into impact-based risk information, rather than in establishing uniformly skilful detection across the study region. Within this framing, the approach offers a means to explore drought risk evolution and potential lead times in regions where soil moisture–vegetation coupling is sufficiently strong. Despite substantial spatial and temporal variability in performance, the framework illustrates how forecast lead time and vegetation response lag can be combined to support impact-oriented drought risk awareness.

Supplementary Information for Chapter 6

SI Figures

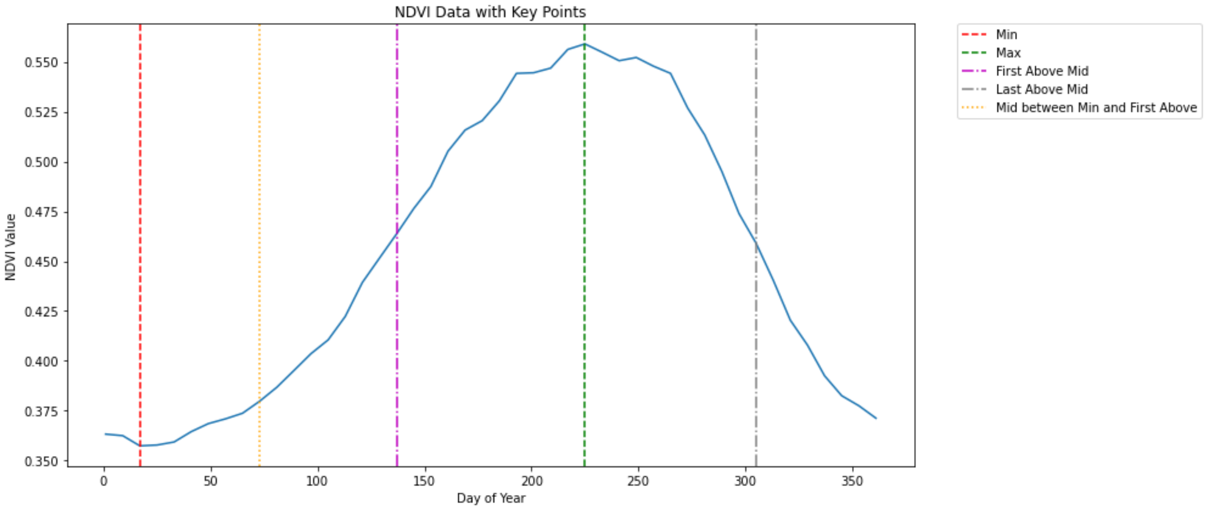


Fig. S6.1. Long-term NDVI phenology averaged over the study area. The annual minimum occurs on January 17; the annual maximum occurs on August 14. NDVI first exceeds the midpoint on May 17 and last remains above it on November 1. Based on this pattern, the growing season is defined as May to October in this study, the primary focus period for VD forecasting.

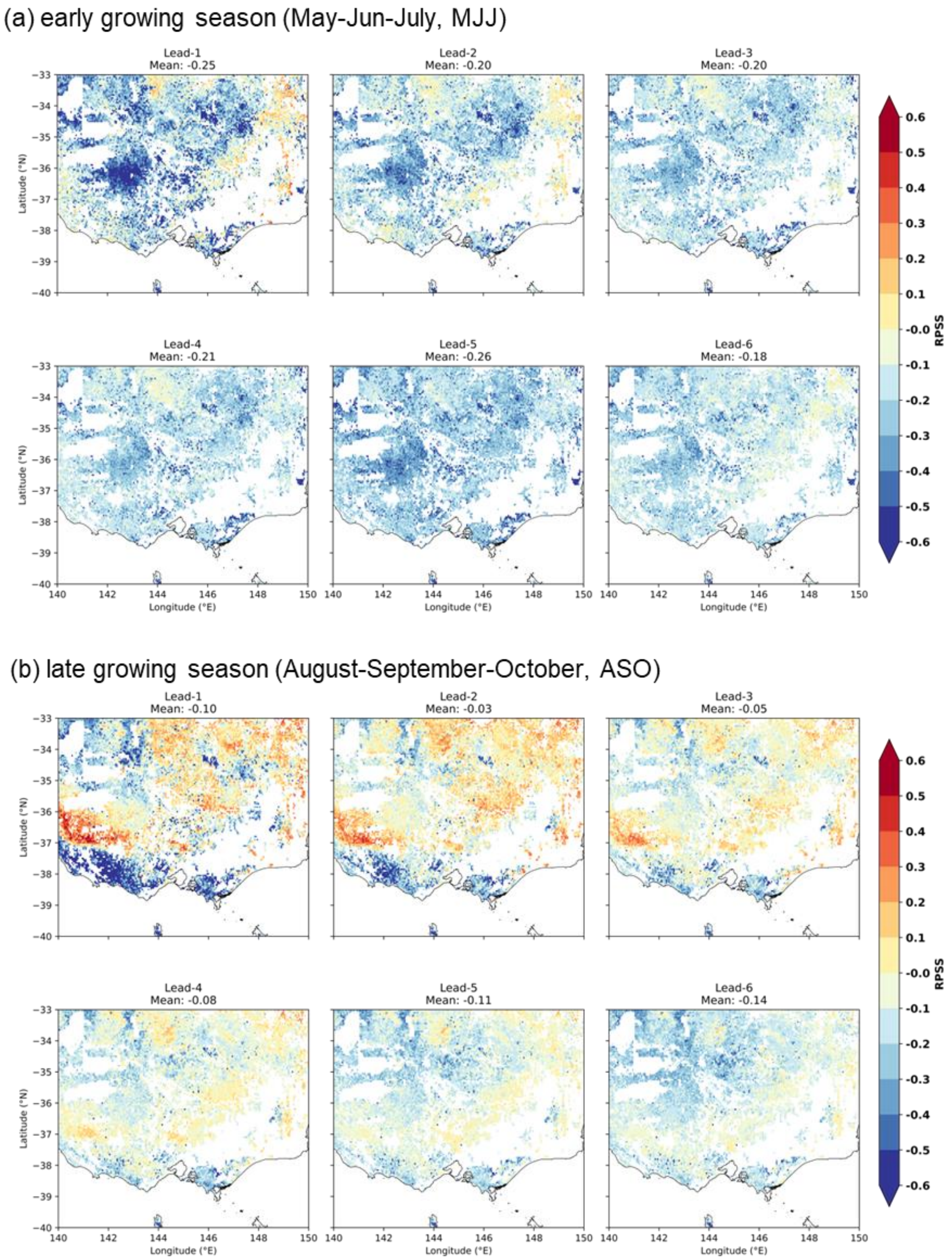


Fig S6.2. Spatial distribution of Ranked Probability Skill Score (RPSS) values for vegetation deterioration (VD) forecasts at 1- to 6-month lead times across southeastern Australia, for the seasonal periods of (a) May–June–July (MJJ) and (b) August–September–October (ASO) during 2001–2017. Warmer colours (redder shades) indicate higher forecast skill.

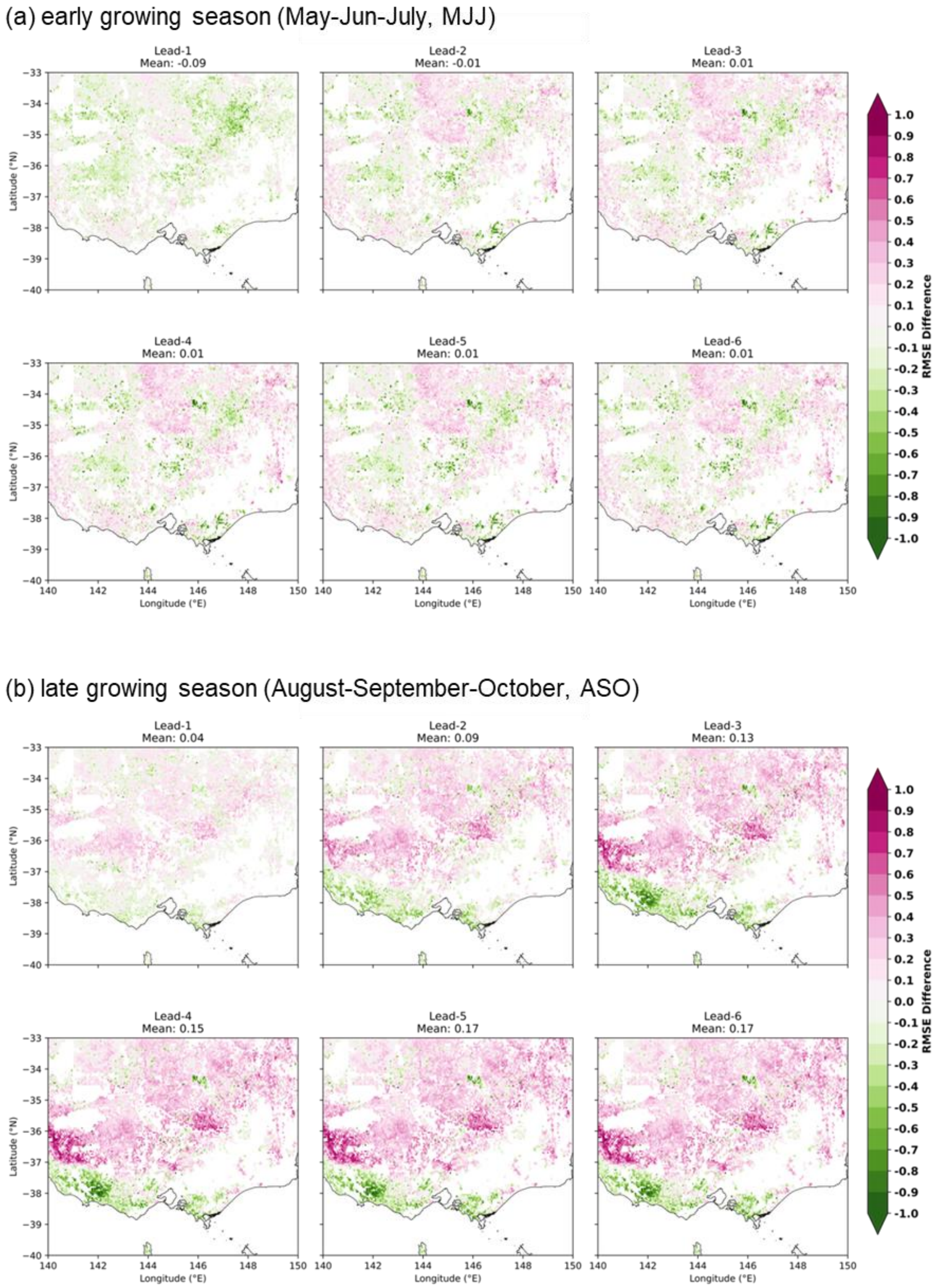


Fig. S6.3. Deterministic forecast skill for vegetation deterioration (VD) events in southeastern Australia, shown separately for (a) the early growing season (May–June–July, MJJ) and (b) the late growing season (August–September–October, ASO). Forecast skill is represented by the difference in root mean square error (RMSE) between hindcasts and the historical baseline (i.e., $\mathrm{RMSE}_{hindcast} - \mathrm{RMSE}_{historical}$) for 1- to 6-month lead times during 2001–2017. The mean RMSE difference, spatially averaged across the study region, is reported above each panel. Negative values (green) indicate improved forecast skill relative to the historical baseline, while positive values (pink) indicate degradation.

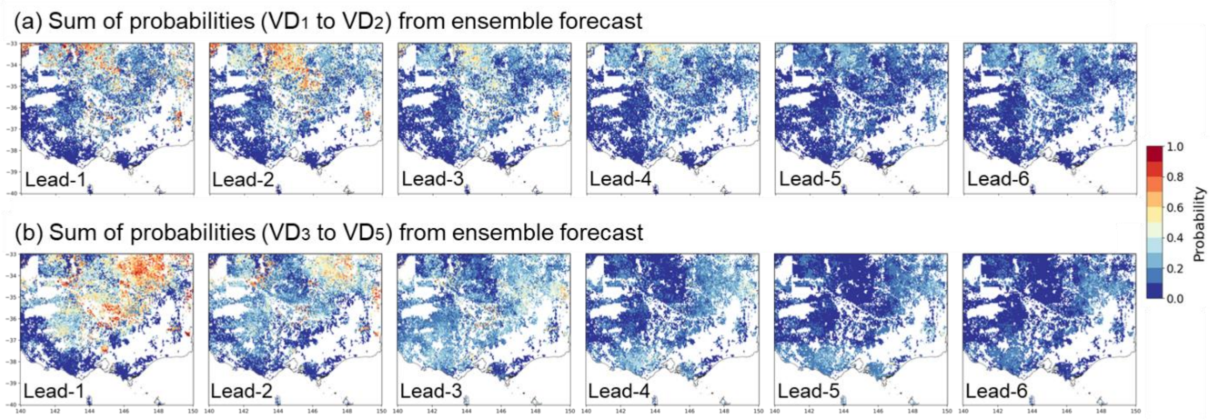


Fig. S6.4. Spatial distribution of forecast vegetation deterioration probabilities for August 2002 across southeastern Australia. (a) Sum of probabilities for mild risk interval (VD_1-VD_2) and (b) severe risk interval (VD_3-VD_5), shown for forecast lead times from 1 to 6 months. Values range from 0 to 1, with higher values indicating greater probability of deterioration occurrence.

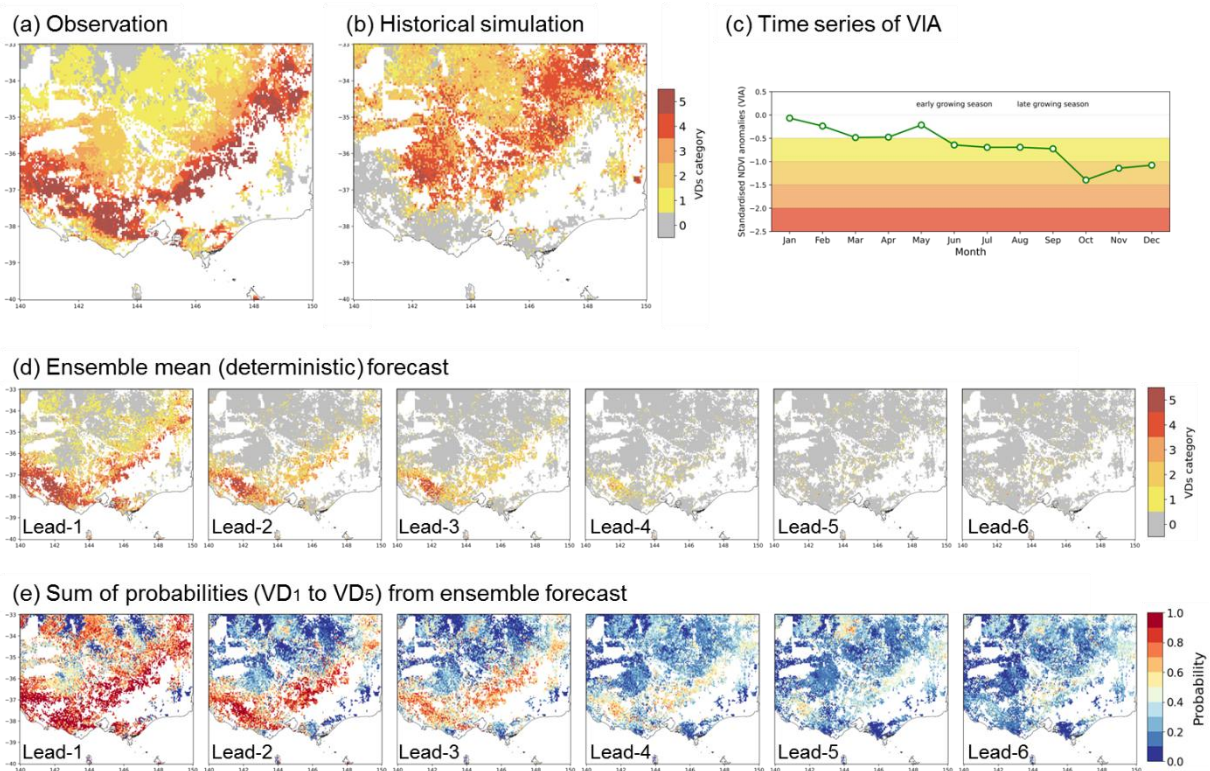


Fig. S6.5. Same as Fig. 6.8 but for October 2006.

SI tables

Table S6.1. Root Mean Square Error (RMSE) of deterministic VD forecasts across different risk intervals and seasons at 1- to 6-month lead times relative to the historical baseline (2001–2017).

Risk interval	Season	Hist	M1	M2	M3	M4	M5	M6
All Risk	ALL	0.97	0.94	1.01	1.04	1.05	1.06	1.06
	MJJ	1	0.91	0.99	1.01	1.01	1.01	1.01
	ASO	0.93	0.97	1.03	1.06	1.09	1.1	1.1
Mild Risk	ALL	0.65	0.67	0.71	0.71	0.71	0.71	0.71
	MJJ	0.67	0.68	0.74	0.74	0.74	0.74	0.74
	ASO	0.61	0.64	0.66	0.67	0.67	0.67	0.67
Severe Risk	ALL	0.85	0.87	0.97	1.05	0.87	0.71	0.71
	MJJ	0.82	0.68	0.6	0.56	0.44	0.6	0.6
	ASO	0.82	0.95	1.02	1.07	0.9	0.73	0.73

Notes: Hist refers to the historical baseline. M1–M6 indicate hindcasts with 1- to 6-month lead times.

Table S6.2. Relative bias (%) of deterministic VD forecasts across different risk intervals and seasons at 1- to 6-month lead times relative to the historical baseline (2001–2017).

Risk interval	Season	Hist	M1	M2	M3	M4	M5	M6
All Risk	ALL	1.2	-20.6	-36.3	-41.5	-43.2	-44.1	-44.1
	MJJ	2.7	-21.0	-37.5	-42.3	-43.1	-43.6	-43.6
	ASO	0.1	-19.3	-33.7	-39.2	-41.9	-43.0	-43.0
Mild Risk	ALL	-10.5	-21.1	-34.6	-39.6	-41.7	-42.8	-42.8
	MJJ	-10.1	-20.8	-35.2	-40.5	-41.6	-42.3	-42.2
	ASO	-8.2	-18.3	-30.9	-35.6	-38.9	-40.6	-40.6
Severe Risk	ALL	31.8	-16.5	-35.5	-40.7	-42.0	-42.5	-42.5
	MJJ	69.8	-6.5	-32.4	-36.9	-37.2	-37.5	-37.5
	ASO	26.3	-13.9	-30.0	-36.3	-38.7	-39.4	-39.4

Notes: Hist refers to the historical baseline. M1–M6 indicate hindcasts with 1- to 6-month lead times.

Table S6.3. Overestimation frequency (%) of deterministic VD forecasts across different risk intervals and seasons at 1- to 6-month lead times relative to the historical baseline (2001–2017).

Risk interval	Season	Hist	M1	M2	M3	M4	M5	M6
All Risk	ALL	24.3	12.9	5.5	2.8	2.0	1.5	1.5
	MJJ	26.8	13.5	5.8	2.6	1.9	1.5	1.5
	ASO	21.9	12.3	5.3	3.0	2.0	1.6	1.6
Mild Risk	ALL	18.6	10.8	4.9	2.7	2.0	1.6	1.6
	MJJ	20.7	11.6	5.5	2.6	1.9	1.5	1.5
	ASO	16.5	10.0	4.4	2.8	2.0	1.6	1.6
Severe Risk	ALL	53.2	36.0	32.4	30.0	31.4	29.7	29.7
	MJJ	63.1	47.9	44.3	29.7	29.9	32.5	32.5
	ASO	48.8	31.2	30.7	29.9	31.8	28.1	28.1

Notes: Hist refers to the historical baseline. M1–M6 indicate hindcasts with 1- to 6-month lead times.

Table S6.4. Underestimation frequency (%) of deterministic VD forecasts across different risk intervals and seasons at 1- to 6-month lead times relative to the historical baseline (2001–2017).

Risk interval	Season	Hist	M1	M2	M3	M4	M5	M6
All Risk	ALL	28.5	36.2	41.0	42.4	42.8	43.1	43.0
	MJJ	30.1	38.6	44.2	45.7	46.0	46.3	46.2
	ASO	26.8	33.8	37.8	39.1	39.6	39.9	39.9
Mild Risk	ALL	26.8	32.6	36.7	37.8	38.1	38.4	38.4
	MJJ	29.2	35.7	40.4	41.7	42.0	42.2	42.2
	ASO	24.0	29.4	32.7	33.7	34.1	34.4	34.4
Severe Risk	ALL	33.0	51.4	57.3	60.6	58.3	54.6	54.7
	MJJ	24.1	36.7	38.4	52.0	38.8	51.7	51.7
	ASO	36.8	58.7	60.5	61.1	60.0	56.6	56.6

Notes: Hist refers to the historical baseline. M1–M6 indicate hindcasts with 1- to 6-month lead times.

Table S6.5. RMSE differences of deterministic vegetation deterioration forecasts across key case study regions at 1- to 6-month lead times, disaggregated by season (ALL, MJJ, ASO) during 2001–2017.

Key region	Season	M1	M2	M3	M4	M5	M6
Riv	ALL	0.01	0.14	0.17	0.18	0.18	0.18
	MJJ	-0.02	0.15	0.18	0.19	0.18	0.18
	ASO	0.04	0.15	0.18	0.19	0.2	0.2
Wim	ALL	0.01	0.15	0.24	0.29	0.3	0.3
	MJJ	-0.06	0.04	0.05	0.05	0.05	0.05
	ASO	0.09	0.27	0.44	0.53	0.54	0.54
GoV	ALL	0.05	0.18	0.25	0.31	0.33	0.32
	MJJ	-0.04	0.1	0.15	0.18	0.18	0.18
	ASO	0.13	0.26	0.34	0.44	0.47	0.47
SwS	ALL	-0.03	0.11	0.19	0.21	0.22	0.22
	MJJ	-0.11	0.07	0.13	0.14	0.14	0.14
	ASO	0.05	0.16	0.25	0.29	0.31	0.31

Notes: M1–M6 indicate hindcasts with 1- to 6-month lead times. Seasons: ALL = May–October (full), MJJ = May–July (early), ASO = August–October (late).

Table S6.6. Ranked Probability Skill Score (RPSS) of probabilistic vegetation deterioration forecasts across key case study regions at 1- to 6-month lead times, disaggregated by season (ALL, MJJ, ASO) during 2001–2017.

Key region	Season	M1	M2	M3	M4	M5	M6
Riv	ALL	-0.01	0	-0.06	-0.07	-0.13	-0.14
	MJJ	-0.1	-0.1	-0.13	-0.12	-0.19	-0.15
	ASO	0.08	0.1	0	-0.02	-0.07	-0.15
Wim	ALL	0.05	0.01	-0.02	-0.07	-0.14	-0.12
	MJJ	-0.15	-0.18	-0.17	-0.15	-0.24	-0.15
	ASO	0.24	0.18	0.11	-0.01	-0.07	-0.1
GoV	ALL	0.04	0.04	-0.02	-0.06	-0.1	-0.07
	MJJ	-0.02	-0.1	-0.14	-0.16	-0.23	-0.1
	ASO	0.09	0.16	0.08	0.04	0	-0.05
SwS	ALL	0.03	0.04	-0.04	-0.08	-0.11	-0.09
	MJJ	-0.03	-0.05	-0.15	-0.18	-0.21	-0.11
	ASO	0.08	0.11	0.05	0.01	-0.03	-0.08

Notes: M1–M6 indicate hindcasts with 1- to 6-month lead times. Seasons: ALL = May–October (full), MJJ = May–July (early), ASO = August–October (late).

Chapter 7

Conclusions and future research directions

7.1. Research Summary

This thesis advances the field of informative drought early warning by improving the forecasting capability of vegetation deterioration (VD) risks in southeastern Australia's dryland agricultural ecosystems. The merit of my proposed approach lies in its ability to predict vegetation growth conditions for the upcoming month or season in advance, with an accompanying level of skill measured by several metrics. My methodology aims to bridge the gap between current soil moisture conditions and future vegetation status. By developing a predictive framework, I can provide valuable insights into the potential impacts of soil moisture deficits on vegetation health. This information is crucial for enhancing drought preparedness and mitigation strategies of natural capital, landscape function, ecosystem services, and farm productivity, potentially benefiting agricultural businesses and agricultural-dependent communities.

First, this thesis investigated the relationship between vegetation dynamics and soil moisture variability, assessing the predictive performance of copula-based joint probability modelling in identifying critical threshold values that trigger multi-categorical VD risks (VD_1-VD_5). Second, the thesis introduced an innovative 'threshold event-matching' framework, which significantly enhances the predictive capability for infrequent severe VD events. Third, by integrating copula-based joint probability modelling with the event-matching framework, the study refined and validated a robust prediction approach capable of capturing VD events across both frequent mild and infrequent severe severities, with particular relevance to agricultural vegetation systems. Finally, this thesis demonstrated the operational applicability of the established framework by applying it to ensemble soil moisture forecast datasets, quantifying both predictive skill and practical early warning horizons. This integration provides critical lead-time information for drought risk management, enhancing the effectiveness of drought early warning systems across the study region and potentially beyond to other Australian agricultural regions.

Four key research questions were determined for the thesis:

RQ1. (Chapter 3): *Can the joint probability distribution approach (e.g., copula-based models) be used to develop a statistical framework for predicting multi-categorical* *vegetation deterioration events based on historical root-zone soil moisture simulations?*

RQ2. (Chapter 4): *Do in-situ soil moisture observations, through a 'threshold event-matching' approach, offer potential capability in predicting infrequent-severe vegetation deterioration events, and if so, what is the lead time?*

RQ3. (Chapter 5): *Can the 'threshold event-matching' approach improve predictions of vegetation deterioration severity at large scales, and how does integrating with copula-based models enhance prediction accuracy for different land use types in southeastern Australia?*

RQ4. (Chapter 6): *To what lead time can ensemble soil moisture forecasts skilfully predict vegetation deterioration risks, and how can deterministic and probabilistic forecasts be applied into drought early warning systems?*

These questions were addressed in corresponding chapters as indicated above in brackets. The following sections summarise and conclude the research presented in this thesis, with Section 7.2 proposing future research directions.

Previous research has demonstrated that vegetation degradation strongly correlates with soil moisture drought stress, particularly in dryland regions (Zhang *et al.,* 2023). Root-zone soil moisture exhibits paramount importance for terrestrial dryland vegetation ecosystems compared to other hydro-meteorological factors (Chen et al., 2014; Akuraju *et al.,* 2017). Consequently, this thesis primarily utilised root-zone soil moisture as the only predictive variable. Chapter 3 developed a prediction framework based on standardised anomalies (i.e., z-scores) of root-zone soil moisture to predict the severity category of VD events, as classified from a satellite-derived vegetation index (i.e., NDVI) in southeastern Australia. The quantitative relationships between soil moisture and vegetation anomalies were established through time-lagged Spearman correlation analyses and copula-based joint probability modelling. Using this probabilistic framework, the conditional probabilities of multi-categorical VD risks were estimated under varying scenarios of root-zone soil moisture anomalies. The analysis highlighted the complex and region-specific nature of vegetation responses to soil moisture deficits, with response times mainly varying from 8 to 40 days across different ecosystems. Although the copula-based prediction framework proved effective in anticipating frequent-mild vegetation drought impacts, it showed limitations in predicting infrequent-severe events. This chapter advances drought early warning systems by quantifying vegetation deterioration risks using soil moisture thresholds across multiple vegetation categories, thereby providing valuable insights for agricultural and environmental management practices.

To enhance the predictive capability for infrequent severe events, Chapter 4 developed an innovative threshold-based event-matching methodology, using in-situ soil moisture observations from the central and western United States of America. This approach proved particularly effective in capturing severe VD events? triggered by soil moisture anomaly signals. In details, the investigation assessed the predictive potential of soil moisture observations for satellite-derived vegetation drought impacts by identifying critical thresholds that delineate the soil moisture stress and VD classifications. The predictive performances were measured by F-score to balance the trade-off between false alarms and missing events across varying soil moisture depths and VD classifications. The results demonstrated that the threshold event-matching approach significantly outperformed random prediction references, as evidenced by the positive improvements in F-scores. Notably, enhanced predictability was achieved for more infrequent-severe VD events (e.g., -2.0 standard deviations), with optimal z-score threshold combinations aligning the similar values of soil moisture anomaly and vegetation index anomaly thresholds. The framework also provided effective lead times of approximately 2-6 weeks, with agricultural regions showing markedly superior predictive capability compared to forest and natural vegetation zones. Overall, the methodology developed in Chapter 4 demonstrated considerable strength in forecasting moderate to extreme VD events during drought seasons, with enhanced predictive capability for drought early warning systems.

Chapter 5 extended the analytical framework established in Chapter 4 by investigating the predictive capability of root-zone soil moisture as simulated by the Australian Water Resource Assessment (AWRA-L) land surface model across extensive spatial domains. This chapter systematically assessed the performance of individual copula-based modelling, event-matching approaches, and their integrated frameworks in predicting VD risks across southeastern Australia. The findings revealed distinct complementary strengths: copula-based prediction frameworks tended to underestimated VD risks, showing a bias toward mild-to-moderate categories (VD_1-VD_2) while underrepresenting severe categories (VD_3-VD_5). In contrast, ‘event-matching’ frameworks showed a tendency toward overestimation with reduced accuracy in mild-to-moderate cases (VD_1-VD_2) and reasonable performance in severe categories (VD_3-VD_5). While neither the standalone copula-based nor ‘event-matching’ models demonstrated consistent superiority over the reference persistence model ($VD_t = VD_{t-1}$) for near-future forecasting, their integration provided valuable complementary insights into VD risks derived from soil moisture signals. The robustness of this integrated approach was further corroborated through a case study of the 2002 drought, in which the combined framework achieved improved predictive accuracy. This methodology proved particularly effective in drought-prone agricultural land use regions while providing notably shorter response scope compared to other land use regions. Ultimately, this chapter contributed significant insights into the operational feasibility of soil moisture-based predictions for drought risk management, particularly in regions characterised by diverse agricultural land-use patterns.

The final research question addressed in Chapter 6 examined the use of seasonal soil moisture forecasts to explore VD risk prediction. Specifically, it assessed the skill of probabilistic and deterministic VD forecasts during the annual growing season (May-October) using ensemble soil moisture hindcasts at lead times ranging from 1 to 6 months. The findings indicated that, although forecast skill decreases with longer lead times, the integrated copula- and event-matching-based framework effectively forecast the spatial-temporal patterns and severity of VD risks across the study area. Rather than demonstrating uniformly robust forecast skill, the analysis illustrated how combining soil moisture forecast lead time with vegetation response lag provides a structured framework for interpreting potential VD risk evolution over subseasonal to seasonal timescales, particularly in regions where soil moisture–vegetation coupling is relatively strong. This horizon combines the inherent vegetation-soil moisture response time (0-2 months) with the extended lead time (1-3 months) provided by soil moisture forecasts. The probabilistic VD risk forecasts provided sub-seasonal to seasonal outlooks, with the probabilistic approach particularly effective in distinguishing between mild and severe VD risks, maintaining prediction confidence above 50% for VD risks up to 4 months across the majority of the study area. The framework proved especially effective in regard to VD early warning capability during the late growing season (August–October), demonstrating higher predictive skill in inland pastoral and dryland cropping zones. Overall, the framework provided reliable and valuable probabilistic insights into the likelihood and severity of drought–induced vegetation stress. These results demonstrate that integrating seasonal soil moisture forecasts into the predictive model developed in Chapter 5 can extend the overall early warning horizon, thereby offering additional lead time for implementing drought preparedness and mitigation strategies. The early warning signals generated by this framework can help in effectively informing decision-making for the agricultural sector of southeastern Australia.

In summary, this thesis presents a comprehensive framework for forecasting vegetation deterioration risks using soil moisture anomalies, with a particular emphasis on the agricultural regions of southeastern Australia. The core studies collectively represent a significant advancement in impact-based drought early warning capabilities through the development and validation of novel predictive frameworks that enhance the forecasting skill across the varying severity levels of VD events. Furthermore, the integration of ensemble soil moisture hindcasts may help extend the temporal reach of drought early warning systems.

7.2. Limitations, knowledge gaps, and future research directions

This research has significantly advanced the field of drought impact early warning through developing a comprehensive categorical VD forecasting framework. Despite these advancements, several limitations remain that warrant further investigation. Future work should focus on enhancing the operational efficiency of the framework, expanding its geographical scope, and integrating additional environmental variables to improve real-time drought predictions. This will provide stakeholders with the tools necessary to mitigate drought impacts on agriculture and the environment, ultimately enhancing drought resilience across Australia. Suggested directions for future research include: (1) refinement of the multiple model integration approach; (2) integration of hydrometeorological and oceanic-atmospheric variables as additional predictors; (3) integration of additional satellite-based vegetation products in drought Impact assessment; (4) expansion of model spatial scale; (5) framework transformation with operational seasonal soil moisture forecasts; (6) Linking economic cost-loss analysis to the VD forecast framework; and (7) Validating the framework in the context of climate change. Each of the areas is discussed below.

(1). Refinement of the multiple model integration approach

This study underscores the need for advancements in the vegetation deterioration (VD) prediction framework to enhance predictive accuracy in regions with persistently low forecast skill, including coastal and other substantial fractions of inland regions (Chapter 6). This can possibly be achieved through region-specific adaptations that emphasise understanding localised vegetation response patterns to hydrometeorological variability. Optimising prediction capabilities requires refining the integration of multiple models, particularly through advanced hybridisation of copula-based and event-matching methods. A key development area involves spatially explicit weighting schemes that allow geographically varying contributions from copula-based and event-matching models based on local performance metrics (Hao *et al.,* 2018), together with dynamic threshold adjustments for evolving climate conditions and regional characteristics (De Kauwe *et al.,* 2020). Additionally, as discussed in Chapter 5, incorporating vegetation condition persistence as the initial VD categories in hybrid models alongside the antecedent soil moisture anomalies may enhance forecast accuracy, particularly for agricultural regions where short-term persistence forecasts show potential for improvement.

(2). Integration of hydrometeorological and oceanic-atmospheric variables as additional predictors

The current framework demonstrates context-dependent utility using root-zone soil moisture as the primary predictor of vegetation health. However, opportunities exist to enhance the framework by incorporating additional environmental parameters, such as hydrometeorological (e.g., maximum temperature extremes, atmospheric evapotranspiration (ET) demand), along with dominant oceanic-atmospheric drivers of intra-seasonal or interannual droughts in Australia, including the El Niño–Southern Oscillation (ENSO), the Indian Ocean Dipole (IOD), the Madden-Julian Oscillation (MJO), and the Southern Annular Mode (SAM) (Feng *et al.,* 2020; Holgate *et al.,* 2020; BoM, 2020). Integrating these parameters into VD predictive models may improve the representation of climate–vegetation linkages relevant for drought forecasting, particularly in regions where vegetation responses are not well explained by soil moisture alone. This further research direction will contribute to more robust VD predict framework, particularly for regions with complex environmental interactions and diverse vegetation ecosystems. Nonetheless, future developments must also account for the sample size limitations associated with higher-dimensional modelling.

(3). Integration of additional satellite-based vegetation products in drought Impact assessment

The current framework can adapt to alternative vegetation condition indicators beyond the currently standalone use of the Normalised Difference Vegetation Index (NDVI). Future research should explore integrating other satellite-derived vegetation products, such as near-infrared reflectance of vegetation (NIRv), solar-induced chlorophyll fluorescence (SIF), kernel NDVI (kNDVI), and vegetation optical depth (VOD). These indices offer complementary perspectives on vegetation functionality, particularly photosynthetic activity and carbon stock dynamics in terrestrial ecosystems under hydrometeorological stress (Badgley *et al.,* 2017; Camps-Valls *et al.,* 2021; Liu *et al.,* 2015; Sun *et al.,* 2017). Expanding to include gross primary production (GPP) and leaf area index (LAI) measurements would further deepen our understanding of ecosystem carbon stock dynamics and structural responses to varying drought stress scenarios (Smith *et al.,* 2019). This methodological expansion would significantly advance our comprehension of vegetation-environment interactions under drought conditions, enhancing the framework's capacity to capture complex ecosystem responses and contributing to more robust drought prediction and management strategies.

(4). Expansion of model spatial scale

While the current research framework has proven effective within southeastern Australia, including Victoria, the Australian Capital Territory, and southern New South Wales, its applicability to other significant agricultural production regions remains unexplored. Future research should focus on expanding this framework across the Australian continental landscape, extending from Queensland's intensive farming zones through South Australia's coastal agricultural districts to Western Australia's diverse agricultural regions. This expansion would encompass both winter- and summer-dominant rainfall regions, representing a broad spectrum of Australian farming systems (Hochman *et al.,* 2017). The validation process should tune and assess the framework performance across these varied landscapes, considering each region's unique agricultural practices and environmental conditions. This expansion would enhance the framework's utility for national agricultural planning and drought risk management, offering broader applicability across diverse ecological and climatic contexts and contributing to more effective continental-scale drought risk management strategies.

(5). Framework transformation with operational seasonal soil moisture forecasts

The shift from hindcast analyses (Chapter 6) to operational forecasting frameworks introduces both challenges and opportunities. The current study used hindcast data at a monthly resolution from 2001 to 2017, while the Australian Water Outlook (AWO) operational system employs a sophisticated 99-member time-lagged ensemble methodology for real-time forecasting. This system, available since October 2021 (https://awo.bom.gov.au/about/overview/forecasts), significantly enhances ensemble dimensionality compared to the 27-member hindcast ensemble in this study. The AWO operational framework provides soil moisture estimations using ensemble percentiles (5th, 25th, 50th, 75th, and 95th), with forecast lead times from one to three months. Future research should integrate these percentile-based soil moisture forecasts into the VD prediction framework. This will require refining statistical techniques to accommodate the utilisation of the full percentile distribution in VD risk prediction. Furthermore, future research could also explore optimisation, potentially incorporating weighted ensemble approaches that account for forecast uncertainty with the increased ensemble number across different lead times and seasons. Integrating these operational soil moisture forecasts represents a critical step towards developing a real-time VD forecasting system for impact-based drought early warning systems.

(6). Linking economic cost-loss analysis to the VD forecast framework

Future research could explore how VD forecasts may be linked to external economic cost–loss or cost–benefit analysis frameworks, allowing forecast information to be interpreted within individual business contexts. Such linkages would enable decision-making to remain with individual business holders, where economic outcomes depend on specific risk preferences, management practices, input costs, and commodity prices, rather than being embedded directly within the forecasting framework. Research could examine the relationship between VD categories and economic metrics, such as income variability (ABBERS, 2022), food security (Rachester, 2023), yield outcomes, ecological services values (Crausbay *et al.,* 2017), and potential loss of wildfire (Jolly *et al.,* 2015), to identify critical thresholds or tipping points for proactive intervention (Klein *et al.,* 2015). Implementing this relative value analysis plot through cost-loss ratio assessments (Busker *et al.,* 2023; Shyrokaya *et al.,* 2023;) align VD forecasts could optimise early action triggers, facilitating the evaluation of intervention timing and resource allocation before drought events, thus enhancing stakeholder decision-making across diverse agricultural and environmental sectors (Busker *et al.,* 2023; Richardson, 2000; Wilks, 2001).

(7). Evaluating the framework in the context of climate change

The research framework is increasingly significant within the context of accelerating climate change, which is expected to intensify the frequency and severity of droughts in southern Australia (Hochman *et al.,* 2017). This underscores the need for reliable drought impact early warning tools for long-term climate adaptation, particularly in regions vulnerable to drought-induced vegetation stress. The Australian Water Outlook's national hydrological projections (https://awo.bom.gov.au/about/overview/projections) provide a temporal framework for assessing future soil moisture conditions, focusing on 2030, 2050, 2070, and 2085 (BoM, 2020). Future research should examine the framework's predictive capability under projected climate scenarios, particularly for RCP4.5 and RCP8.5 emission scenarios. This will involve adapting soil moisture-vegetation response relationships to account for enhanced atmospheric CO_2 concentrations and altered precipitation patterns, as well as refining prediction frameworks to accommodate changing temperature extremes and soil moisture availability (Knauer *et al.,* 2017). Extending and evaluating this framework under projected climate scenarios is crucial for improving drought preparedness and mitigation strategies, particularly in agricultural regions most vulnerable to climate change impacts. This research direction will contribute to adaptive management approaches that consider both current and future climate scenarios in drought risk assessment and mitigation planning.

Bibliography

A

AghaKouchak, A., Farahmand, A., Melton, F.S., Teixeira, J., Anderson, M.C., Wardlow, B.D., Hain, C.R. (2015). Remote sensing of drought: Progress, challenges and opportunities: remote sensing of drought. *Reviews of Geophysics*, 53, 452–480. https://doi.org/10.1002/2014RG000456

Ahmed, M., Else, B., Eklundh, L., Ardö, J., Seaquist, J. (2017). Dynamic response of NDVI to soil moisture variations during different hydrological regimes in the Sahel region. *International Journal of Remote Sensing,* 38, 5408–5429. https://doi.org/10.1080/01431161.2017.1339920

Akuraju, V. R., Ryu, D., George, B., Ryu, Y., & Dassanayake, K. (2017). Seasonal and inter-annual variability of soil moisture stress function in dryland wheat field, Australia. *Agricultural and Forest Meteorology*, *232*, 489-499. https://doi.org/10.1016/j.agrformet.2016.10.007

Anderegg, W. R. L., Trugman, A. T., Badgley, G., Konings, A. G., & Shaw, J. D. (2020). Divergent forest sensitivity to repeated extreme droughts. *Nature Climate Change*, 10(12), 1091–1095. https://doi.org/10.1038/s41558-020-00919-1

Arnal, L., Cloke, H.L., Stephens, E., Wetterhall, F., Prudhomme, C., Neumann, J., Krzeminski, B., Pappenberger, F. (2018). Skilful seasonal forecasts of streamflow over Europe? *Hydrology and Earth System Sciences,* 22, 2057–2072. https://doi.org/10.5194/hess-22-2057-2018

Arora, V.K., Boer, G.J. (2003). A Representation of Variable Root Distribution in Dynamic Vegetation Models. *Earth Interactions,* 7, 1–19. https://doi.org/10.1175/1087-3562(2003)007<0001:AROVRD>2.0.CO;2

Asoka, A., Mishra, V. (2015). Prediction of vegetation anomalies to improve food security and water management in India. *Geophysical Research Letters,* 42, 5290–5298. https://doi.org/10.1002/2015GL063991

Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES) Australian Crop Report No. 197 February 2021. Australian Government Department of Agriculture, Water and the Environment; Canberra, Australia: 2021. https://doi.org/10.25814/xqy3-sx57.

B

Bachmair, S., Kohn, I., Stahl, K. (2015). Exploring the link between drought indicators and impacts. *Natural Hazards and Earth System Sciences,* 15, 1381–1397. https://doi.org/10.5194/nhess-15-1381-2015

Bachmair, S., Svensson, C., Hannaford, J., Barker, L.J., Stahl, K. (2016a). A quantitative analysis to objectively appraise drought indicators and model drought impacts. *Hydrology and Earth System Sciences,* 20, 2589–2609. https://doi.org/10.5194/hess-

20-2589-2016

Bachmair, Sophie, Stahl, K., Collins, K., Hannaford, J., Acreman, M., Svoboda, M., Knutson, C., Smith, K.H., Wall, N., Fuchs, B., Crossman, N.D., Overton, I.C. (2016b). Drought indicators revisited: the need for a wider consideration of environment and society: Drought indicators revisited. *Wiley Interdisciplinary Reviews: Water,* 3, 516–536. https://doi.org/10.1002/wat2.1154

Badgley, G., Field, C.B., Berry, J.A. (2017). Canopy near-infrared reflectance and terrestrial photosynthesis. *Science Advances*, 3, e1602244. https://doi.org/10.1126/sciadv.1602244

Baldocchi, D.D. (2020). How eddy covariance flux measurements have contributed to our understanding of Global Change Biology. *Global Change Biology*, 26, 242–260. https://doi.org/10.1111/gcb.14807

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*, 111886. https://doi.org/10.1016/j.rse.2020.111886

Bell, J.E., Palecki, M.A., Baker, C.B., Collins, W.G., Lawrimore, J.H., Leeper, R.D., Hall, M.E., Kochendorfer, J., Meyers, T.P., Wilson, T., Diamond, H.J. (2013). U.S. Climate Reference Network Soil Moisture and Temperature Observations. *Journal of Hydrometeorology,* 14, 977–988. https://doi.org/10.1175/JHM-D-12-0146.1

Bento, V.A., Gouveia, C.M., DaCamara, C.C., Trigo, I.F. (2018). A climatological assessment of drought impact on vegetation health index. *Agricultural and forest meteorology,* 259, 286–295. https://doi.org/10.1016/j.agrformet.2018.05.014

Blauhut, V., Gudmundsson, L., Stahl, K. (2015). Towards pan-European drought risk maps: quantifying the link between drought indices and reported drought impacts. *Environment Research Letters*, 10, 014008. https://doi.org/10.1088/1748-9326/10/1/014008

Boken, V.K., Cracknell, A.P., Heathcote, R.L. (2005). Monitoring and predicting agricultural drought: a global study. New York (USA) *Oxford University Press.* https://doi.org/10.1093/oso/9780195162349.001.0001

Bolten, J.D., Crow, W.T. (2012). Improved prediction of quasi-global vegetation conditions using remotely-sensed surface soil moisture*. Geophysical Research Letters,* 39. https://doi.org/10.1029/2012GL053470

Brock, F.V., Crawford, K.C., Elliott, R.L., Cuperus, G.W., Stadler, S.J., Johnson, H.L., Eilts, M.D. (1995). The Oklahoma Mesonet: A Technical Overview. *Journal of Atmospheric and Oceanic Technology*, 12, 5–19. https://doi.org/10.1175/1520-0426(1995)012<0005:TOMATO>2.0.CO;2

Broich, M., Huete, A., Paget, M., Ma, X., Tulbure, M., Coupe, N.R., Evans, B., Beringer, J., Devadas, R., Davies, K., Held, A. (2015). A spatially explicit land surface phenology data product for science, monitoring and natural resources management applications. *Environmental Modelling & Software,* 64, 191–204. https://doi.org/10.1016/j.envsoft.2014.11.017

Bruno Soares, M., Daly, M., & Dessai, S. (2018). Assessing the value of seasonal climate forecasts for decision-making. *Wiley Interdisciplinary Reviews: Climate Change*, *9*(4), e523. https://doi.org/10.1002/wcc.523

Brust, C., Kimball, J.S., Maneta, M.P., Jencso, K., Reichle, R.H. (2021). DroughtCast: A

Machine Learning Forecast of the United States Drought Monitor. *Frontiers in Big Data*, 4, 773478. https://doi.org/10.3389/fdata.2021.773478

C

Camps-Valls, G., Campos-Taberner, M., Moreno-Martínez, Á., Walther, S., Duveiller, G., Cescatti, A., ... & Running, S. W. (2021). A unified vegetation index for quantifying the terrestrial biosphere. *Science Advances*, *7*(9), eabc7447. https://doi.org/10.1126/sciadv.abc7447

Carter, J.O., Hall, W.B., Brook, K.D, McKeon, G.M., Day, K.A. and Paull, C.J. (2000). AussieGRASS: Australian grassland and rangeland assessment by spatial simulation. In: 'Applications of Seasonal Climate Forecasting in Agricultural and Natural Ecosystems – the Australian Experience' (Eds. G. Hammer, N. Nicholls and C. Mitchell). Kluwer Academic Press: Netherlands pp. 329-250 https://link.springer.com/chapter/10.1007/978-94-015-9351-9_20

Chatterjee, S., Desai, A.R., Zhu, J., Townsend, P.A., Huang, J. (2022). Soil moisture as an essential component for delineating and forecasting agricultural rather than meteorological drought. *Remote Sensing of Environment*, 269, 112833. https://doi.org/10.1016/j.rse.2021.112833

Chen, T., de Jeu, R.A.M., Liu, Y.Y., van der Werf, G.R., Dolman, A.J. (2014). Using satellite based soil moisture to quantify the water driven variability in NDVI: A case study over mainland Australia. *Remote Sensing of Environment*, 140, 330–338. https://doi.org/10.1016/j.rse.2013.08.022

Chen, L., & Guo, S. (2019). Copulas and its application in hydrology and water resources (pp. 13-38). *Springer Singapore*. https://doi.org/10.1007/978-981-13-0574-0

Chenu, K., Porter, J. R., Martre, P., Basso, B., Chapman, S. C., Ewert, F., ... & Asseng, S. (2017). Contribution of crop models to adaptation in wheat. *Trends in plant science*, 22(6), 472-490. http://dx.doi.org/10.1016/j.tplants.2017.02.003

Coughlan de Perez, E., Harrison, L., Berse, K., Easton-Calabria, E., Marunye, J., Marake, M., Murshed, S.B., Shampa, Zauisomue, E.-H. (2022). Adapting to climate change through anticipatory action: The potential use of weather-based early warnings. *Weather and Climate Extremes*, 38, 100508. https://doi.org/10.1016/j.wace.2022.100508

Coughlan de Perez, E., van den Hurk, B., van Aalst, M.K., Jongman, B., Klose, T., Suarez, P. (2015). Forecast-based financing: an approach for catalyzing humanitarian action based on extreme weather and climate forecasts. *Natural Hazards and Earth System Sciences,* 15, 895–904. https://doi.org/10.5194/nhess-15-895-2015

Crausbay, S. D., Ramirez, A. R., Carter, S. L., Cross, M. S., Hall, K. R., Bathke, D. J., ... & Wilhelmi, O. V. (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

Crow, W.T., Kumar, S.V., Bolten, J.D. (2012). On the utility of land surface models for agricultural drought monitoring. *Hydrology and Earth System Sciences,* 16, 3451–3460. https://doi.org/10.5194/hess-16-3451-2012

Cunha, A. P. M., Alvalá, R. C., Nobre, C. A., & Carvalho, M. A. (2015). Monitoring vegetative drought dynamics in the Brazilian semiarid region*. Agricultural and forest meteorology*, 214, 494-505. https://doi.org/10.1016/j.agrformet.2015.09.010

D

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

Damm, A., Paul-Limoges, E., Haghighi, E., Simmer, C., Morsdorf, F., Schneider, F.D., van der Tol, C., Migliavacca, M. and Rascher, U. (2018). Remote sensing of plant-water relations: An overview and future perspectives. *Journal of plant physiology*, 227, pp.3-19. https://doi.org/10.1016/j.jplph.2018.04.012

De Kauwe, M. G., Medlyn, B. E., Ukkola, A. M., Mu, M., Sabot, M. E., Pitman, A. J., ... & Briggs, P. R. (2020). Identifying areas at risk of drought-induced tree mortality across South-Eastern Australia. *Global Change Biology*, *26*(10), 5716-5733. https://doi.org/10.1111/gcb.15215

Deng, Y., Wang, X., Wang, K., Ciais, P., Tang, S., Jin, L., Li, L., Piao, S. (2021). Responses of vegetation greenness and carbon cycle to extreme droughts in China. *Agricultural and forest meteorology*. 298–299, 108307. https://doi.org/10.1016/j.agrformet.2020.108307

Dey, R., Lewis, S. C., Arblaster, J. M., & Abram, N. J. (2019). A review of past and projected changes in Australia's rainfall. *Wiley Interdisciplinary Reviews: Climate Change*, 10(3), e577. https://doi.org/10.1002/wcc.577

Dirmeyer, P.A., Wu, J., Norton, H.E., Dorigo, W.A., Quiring, S.M., Ford, T.W., Santanello, J.A., Bosilovich, M.G., Ek, M.B., Koster, R.D., Balsamo, G., Lawrence, D.M. (2016). Confronting Weather and Climate Models with Observational Data from Soil Moisture Networks over the United States. *Journal of Hydrometeorology*, 17, 1049–1067. https://doi.org/10.1175/JHM-D-15-0196.1

Donohue, R. J., McVicar, T. R., & Roderick, M. L. (2009). Climate-related trends in Australian vegetation cover as inferred from satellite observations, 1981–2006. *Global Change Biology*, 15(4), 1025–1039. https://doi.org/10.1111/j.1365-2486.2008.01746.x

Dorigo, W. A., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., ... & De Jeu, R. (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

Durante, F., & Sempi, C. (2015). Principles of Copula Theory (1st ed.). *Chapman and Hall/CRC*. https://doi.org/10.1201/b18674

F

Fang, W., Huang, S., Huang, Q., Huang, G., Wang, H., Leng, G., Wang, L., Li, P., Ma, L. (2019). Bivariate probabilistic quantification of drought impacts on terrestrial vegetation dynamics in mainland China. *Journal of Hydrology*, 577, 123980. https://doi.org/10.1016/j.jhydrol.2019.123980

FAO, (2021). The impact of disasters and crises on agriculture and food security: 2021. FAO, Rome, Italy. https://doi.org/10.4060/cb3673en

Farooq, M., Wahid, A., Kobayashi, N., Fujita, D., & Basra, S. M. A. (2009). Plant drought stress: effects, mechanisms and management. In Sustainable agriculture (pp. 153-188). Dordrecht: *Springer Netherlands*. https://doi.org/10.1051/agro:2008021

Farooq, M., Hussain, M., Wahid, A., Siddique, K.H.M. (2012). Drought Stress in Plants: An Overview, in: Aroca, R. (Ed.), Plant Responses to Drought Stress: From Morphological to Molecular Features. *Springer*, Berlin, Heidelberg, pp. 1–33. https://doi.org/10.1007/978-3-642-32653-0_1

Farahmand, A., & AghaKouchak, A. (2015). A generalized framework for deriving nonparametric standardized drought indicators. *Advances in Water Resources*, 76, 140-145. https://doi.org/10.1016/j.advwatres.2014.11.012

Feng, P., Wang, B., Liu, D.L., Waters, C., Yu, Q. (2019). Incorporating machine learning with biophysical model can improve the evaluation of climate extremes impacts on wheat yield in south-eastern Australia. *Agricultural and forest meteorology*. 275, 100–113. https://doi.org/10.1016/j.agrformet.2019.05.018

Feng, P., Wang, B., Luo, J. J., Li Liu, D., Waters, C., Ji, F., ... & Yu, Q. (2020). Using large-scale climate drivers to forecast meteorological drought condition in growing season across the Australian wheatbelt. *Science of the Total Environment*, 724, 138162. https://doi.org/10.1016/j.scitotenv.2020.138162

Flexas, J., Briantais, J.-M., Cerovic, Z., Medrano, H., Moya, I. (2000). Steady-State and Maximum Chlorophyll Fluorescence Responses to Water Stress in Grapevine Leaves: A New Remote Sensing System. *Remote Sensing of Environment,* 73, 283–297. https://doi.org/10.1016/S0034-4257(00)00104-8

Frankenberg, C., Fisher, J.B., Worden, J., Badgley, G., Saatchi, S.S., Lee, J.-E., Toon, G.C., Butz, A., Jung, M., Kuze, A., Yokota, T. (2011). New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophysical Research Letters*, 38. https://doi.org/10.1029/2011GL048738

[dataset] Friedl, Mark, Sulla-Menashe, Damien. (2019). MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006. https://doi.org/10.5067/MODIS/MCD12Q1.006

Frost, A. J., Ramchurn, A., & Smith, A. (2018). The Bureau's operational AWRA-L version 6 model: Technical description and performance. Bureau of Meteorology, Australia. https://awo.bom.gov.au/assets/notes/publications/AWRALv6_Model_Description_Report.pdf

Funk, C., Shukla, S. (2020). Drought Early Warning and Forecasting: Theory and Practice. Elsevier. https://doi.org/10.1016/C2016-0-04328-0

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, 1011–1027. https://doi.org/10.1175/BAMS-D-17-0233.1

G

Getirana, A., Rodell, M., Kumar, S., Beaudoing, H. K., Arsenault, K., Zaitchik, B., ... & Bettadpur, S. (2020). GRACE improves seasonal groundwater forecast initialization over the United States. *Journal of hydrometeorology*, *21*(1), 59-71. https://doi.org/10.1175/JHM-D-19-0096.1

Ghannam, K., Nakai, T., Paschalis, A., Oishi, C.A., Kotani, A., Igarashi, Y., Kumagai, T., Katul, G.G. (2016). Persistence and memory timescales in root-zone soil moisture dynamics. *Water Resources Research,* 52, 1427–1445. https://doi.org/10.1002/2015WR017983

Griffiths M, Smith P, Yan H, Spillman C, Young G, Hudson D. (2023). ACCESS-S2: Updates and improvements to postprocessing pipeline Bureau Research Report, No. 082, Bureau of Meteorology Australia. http://www.bom.gov.au/research/publications/researchreports/BRR-082.pdf

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. https://doi.org/10.1029/2008GL035772

H

Han, Z., Huang, S., Peng, J., Li, J., Leng, G., Huang, Q., Zhao, J., Yang, F., He, P., Meng, X., Li, Z. 2023. GRACE-based dynamic assessment of hydrological drought trigger thresholds induced by meteorological drought and possible driving mechanisms. *Remote Sensing of Environment,* 298, 113831. https://doi.org/10.1016/j.rse.2023.113831

Hao, Z. and AghaKouchak, A. 2013. Multivariate standardized drought index: a parametric multi-index model. *Advances in Water Resources*, 57, pp.12-18. https://doi.org/10.1016/j.advwatres.2013.03.009

Hao, Z., Singh, V.P. 2015. Drought characterization from a multivariate perspective: A review. *Journal of Hydrology*, 527, 668–678. https://doi.org/10.1016/j.jhydrol.2015.05.031

Hao, Z., Hao, F., Singh, V.P., Xia, Y., Ouyang, W., Shen, X. (2016). A theoretical drought classification method for the multivariate drought index based on distribution properties of standardized drought indices. *Advances in Water Resources,* 92, 240–247. https://doi.org/10.1016/j.advwatres.2016.04.010

Hao, Z., Singh, V.P., Xia, Y. (2018). Seasonal Drought Prediction: Advances, Challenges, and Future Prospects. *Reviews of Geophysics*, 56, 108–141. https://doi.org/10.1002/2016RG000549

Hao, Z., Yuan, X., Xia, Y., Hao, F., Singh, V.P. (2017a). An Overview of Drought Monitoring and Prediction Systems at Regional and Global Scales. *Bulletin of the American Meteorological Society,* 98, 1879–1896. https://doi.org/10.1175/BAMS-D-15-00149.1

Hao, Z., Xia, Y., Luo, L., Singh, V. P., Ouyang, W., & Hao, F. (2017b). Toward a categorical drought prediction system based on US Drought Monitor (USDM) and climate forecast. *Journal of Hydrology*, *551*, 300-305. https://doi.org/10.1016/j.jhydrol.2017.06.005

Hao, Z., Hao, F., Singh, V. P., Xia, Y., Shi, C., & Zhang, X. (2018). A multivariate approach for statistical assessments of compound extremes. *Journal of hydrology*, 565, 87-94. https://doi.org/10.1016/j.jhydrol.2018.08.025

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, 485–488. https://doi.org/10.1175/2010BAMS3103.1

Hazra, A., McNally, A., Slinski, K., Arsenault, K.R., Shukla, S., Getirana, A., Jacob, J.P., Sarmiento, D.P., Peters-Lidard, C., Kumar, S.V., Koster, R.D. (2023). NASA's NMME-based S2S hydrologic forecast system for food insecurity early warning in southern Africa. *Journal of hydrology,* 617, 129005. https://doi.org/10.1016/j.jhydrol.2022.129005

Hochman, Z., Gobbett, D. L., & Horan, H. (2017). Climate trends account for stalled wheat yields in Australia since 1990. *Global Change Biology*, 23(5), 2071–2081. https://doi.org/10.1111/gcb.13604

Hoerling, M., Eischeid, J., Kumar, A., Leung, R., Mariotti, A., Mo, K., Schubert, S., Seager, R. (2014). Causes and Predictability of the 2012 Great Plains Drought. *Bulletin of the American Meteorological Society,* 95, 269–282. https://doi.org/10.1175/BAMS-D-13-00055.1

Holgate, C.M., De Jeu, R.A.M., van Dijk, A.I.J.M., Liu, Y.Y., Renzullo, L.J., Vinodkumar, Dharssi, I., Parinussa, R.M., Van Der Schalie, R., Gevaert, A., Walker, J., McJannet, D., Cleverly, J., Haverd, V., Trudinger, C.M., Briggs, P.R. (2016). Comparison of remotely sensed and modelled soil moisture data sets across Australia. *Remote Sensing of Environment,* 186, 479–500. https://doi.org/10.1016/j.rse.2016.09.015

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), e2020GL090238. https://doi.org/10.1029/2020GL090238

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote sensing of environment*, *83*(1-2), 195-213. https://doi.org/10.1016/S0034-4257(02)00096-2

I

Intergovernmental Panel on Climate Change. (2015). Climate Change 2014: Mitigation of Climate Change: Working Group III Contribution to the IPCC Fifth Assessment Report. *Cambridge University Press*, Cambridge. https://doi.org/10.1017/CBO9781107415416

J

Ji, L., Peters, A.J. (2003). Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices. *Remote Sensing of Environment,* 87, 85–98. https://doi.org/10.1016/S0034-4257(03)00174-3

Johnson, F., & Sharma, A. (2015). What are the impacts of bias correction on future drought projections? *Journal of Hydrology*, 525, 472-485. https://doi.org/10.1016/j.jhydrol.2015.04.002

Johnson, F., Stephens, C., & Krogh, M. (2024). Considerations in designing climate change assessments for complex, non-linear hydrological systems. *Journal of Hydrology*, *645*, 132182. https://doi.org/10.1016/j.jhydrol.2024.132182

Joiner, J., Yoshida, Y., Anderson, M., Holmes, T., Hain, C., Reichle, R., Koster, R., Middleton, E., Zeng, F.-W. (2018). Global relationships among traditional reflectance vegetation indices (NDVI and NDII), evapotranspiration (ET), and soil moisture variability on weekly timescales. *Remote Sensing of Environment,* 219, 339–352. https://doi.org/10.1016/j.rse.2018.10.020

Jolly, W. M., Cochrane, M. A., Freeborn, P. H., Holden, Z. A., Brown, T. J., Williamson, G. J., & Bowman, D. M. J. S. (2015). Climate-induced variations in global wildfire danger from 1979 to 2013. *Nature Communications*, 6, 7537. https://doi.org/10.1038/ncomms8537

Jones, D. A., Wang, W., & Fawcett, R. (2009). High-quality spatial climate data-sets for Australia. *Australian Meteorological and Oceanographic Journal*, *58*(4), 233. https://doi.org/10.1071/ES09032

K

Kelman, I., Glantz, M.H. (2014). Early Warning Systems Defined, in: Singh, A., Zommers, Z. (Eds.), Reducing Disaster: Early Warning Systems For Climate Change. *Springer Netherlands,* Dordrecht, pp. 89–108. https://doi.org/10.1007/978-94-017-8598-3_5

Klein, R. J., Midgley, G., Preston, B. L., Alam, M., Berkhout, F., Dow, K., & Shaw, M. R. (2015). Adaptation opportunities, constraints, and limits. In Climate change 2014: Impacts, adaptation, and vulnerability. Part A: Global and sectoral aspects. Contribution of working group II to the fifth assessment report of the intergovernmental panel on climate change (p. 899). Cambridge: Cambridge University Press. In: *IPCC AR5 WGII Chapter 16*. https://www.ipcc.ch/report/ar5/wg2/

Knauer, J., Zaehle, S., Reichstein, M., Medlyn, B. E., Forkel, M., Hagemann, S., & Werner, C. (2017). The response of ecosystem water-use efficiency to rising atmospheric CO_2 concentrations: Sensitivity and large-scale biogeochemical implications. *New Phytologist,* 213(4), 1654-1666. https://doi.org/10.1111/nph.14288

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., Goldberg, M., Schott, T., Guo, W. (2015). Suomi NPP/VIIRS: improving drought watch, crop loss prediction, and food security. *International Journal of Remote Sensing,* 36, 5373–5383. https://doi.org/10.1080/01431161.2015.1095370

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

Konings, A. G., Williams, A. P., & Gentine, P. (2017). Sensitivity of grassland productivity to aridity controlled by stomatal and xylem regulation. *Nature Geoscience*, *10*(4), 284-288. https://doi.org/10.1038/ngeo2903

Konings, A. G., Saatchi, S. S., Frankenberg, C., Keller, M., Quetin, G. R., & Yu, Y. (2021). Detecting forest response to droughts with global observations of vegetation water content. *Global Change Biology,* 27(23), 6026–6039. https://doi.org/10.1111/gcb.15853

Krishnamurthy R, P.K., Fisher, J.B., Choularton, R.J., Kareiva, P.M. (2022). Anticipating drought-related food security changes. *Nature Sustainability*. 1–9. https://doi.org/10.1038/s41893-022-00962-0

L

Laimighofer, J., Laaha, G. (2022). How standard are standardized drought indices? Uncertainty components for the SPI & SPEI case. *Journal of Hydrology,* 613, 128385. https://doi.org/10.1016/j.jhydrol.2022.128385

Li, W., Pacheco-Labrador, J., Migliavacca, M. *et al.* Widespread and complex drought effects on vegetation physiology inferred from space. *Nature Communications,* 14, 4640 (2023). https://doi.org/10.1038/s41467-023-40226-9

Li, X., Piao, S., Huntingford, C., Peñuelas, J., Yang, H., Xu, H., Chen, A., Friedlingstein, P., Keenan, T.F., Sitch, S., Wang, X., Zscheischler, J., Mahecha, M.D. (2023). Global variations in critical drought thresholds that impact vegetation. *National Science Review,* 10, nwad049. https://doi.org/10.1093/nsr/nwad049

Li, Y., van Dijk, A.I.J.M., Tian, S., Renzullo, L.J. (2023). Skill and lead time of vegetation drought impact forecasts based on soil moisture observations. *Journal of Hydrology,* 620, 129420. https://doi.org/10.1016/j.jhydrol.2023.129420

Liu, L., Teng, Y., Wu, J., Zhao, W., Liu, S., Shen, Q. (2020). Soil water deficit promotes the effect of atmospheric water deficit on solar-induced chlorophyll fluorescence. *Science of The Total Environment,* 720, 137408. https://doi.org/10.1016/j.scitotenv.2020.137408

Liu, H., Liu, Y., Chen, Yu, Fan, M., Chen, Yin, Gang, C., You, Y., Wang, Z. (2023). Dynamics of global dryland vegetation were more sensitive to soil moisture: Evidence from multiple vegetation indices. *Agricultural and forest meteorology*, 331, 109327. https://doi.org/10.1016/j.agrformet.2023.109327

Liu, Y. Y., De Jeu, R. A., McCabe, M. F., Evans, J. P., & Van Dijk, A. I. (2011). Global long-term passive microwave satellite-based retrievals of vegetation optical depth. *Geophysical Research Letters*, *38*(18). https://doi.org/10.1029/2011gl048684

Liu, Y.Y., van Dijk, A.I.J.M., de Jeu, R.A.M., Canadell, J.G., McCabe, M.F., Evans, J.P., Wang, G. (2015). Recent reversal in loss of global terrestrial biomass. *Nature Climate Change*, 5, 470–474. https://doi.org/10.1038/nclimate2581

M

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

McDowell, N.G., Ryan, M.G., Zeppel, M.J.B., Tissue, D.T. (2013). Feature: Improving our knowledge of drought-induced forest mortality through experiments, observations, and modeling. *New Phytologist*. 200, 289–293. https://doi.org/10.1111/nph.12502

McKee, T. B., Doesken, N. J., & Kleist, J. (1993, January). 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). https://www.droughtmanagement.info/literature/AMS_Relationship_Drought_Frequency_Duration_Time_Scales_1993.pdf

McPherson, R.A., Fiebrich, C.A., Crawford, K.C., Kilby, J.R., Grimsley, D.L., Martinez, J.E., Basara, J.B., Illston, B.G., Morris, D.A., Kloesel, K.A., Melvin, A.D., Shrivastava, H., Wolfinbarger, J.M., Bostic, J.P., Demko, D.B., Elliott, R.L., Stadler, S.J., Carlson, J.D., Sutherland, A.J. (2007). Statewide Monitoring of the Mesoscale Environment: A Technical Update on the Oklahoma Mesonet. *Journal of Atmospheric and Oceanic Technology,* 24, 301–321. https://doi.org/10.1175/JTECH1976.1

Méndez-Barroso, L.A., Vivoni, E.R., Watts, C.J., Rodríguez, J.C. (2009). Seasonal and interannual relations between precipitation, surface soil moisture and vegetation dynamics in the North American monsoon region. *Journal of Hydrology,* 377, 59–70. https://doi.org/10.1016/j.jhydrol.2009.08.009

Mendicino, G., Senatore, A., Versace, P. (2008). A Groundwater Resource Index (GRI) for drought monitoring and forecasting in a mediterranean climate. *Journal of Hydrology,* 357, 282–302. https://doi.org/10.1016/j.jhydrol.2008.05.005

Meroni, M., Fasbender, D., Rembold, F., Atzberger, C., Klisch, A. (2019). Near real-time vegetation anomaly detection with MODIS NDVI: Timeliness vs. accuracy and effect of anomaly computation options. *Remote Sensing of Environment*, 221, 508–521. https://doi.org/10.1016/j.rse.2018.11.041

Miguez-Macho, G., Fan, Y. (2021). Spatiotemporal origin of soil water taken up by vegetation. *Nature,* 598, 624–628. https://doi.org/10.1038/s41586-021-03958-6

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

Mishra, S., & Datta-Gupta, A. (2017). Applied statistical modeling and data analytics: A practical guide for the petroleum geosciences. *Elsevier.* https://doi.org/10.1016/C2014-0-03954-8

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N. (2021). ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data,* 13, 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

N

Nelsen, R. B. (2006). *An introduction to copulas*. New York, NY: Springer New York. https://doi.org/10.1007/0-387-28678-0_5

Nicolai-Shaw, N., Zscheischler, J., Hirschi, M., Gudmundsson, L., Seneviratne, S.I. (2017). A drought event composite analysis using satellite remote-sensing based soil moisture. *Remote Sensing of Environment,* 203, 216–225. https://doi.org/10.1016/j.rse.2017.06.014

Nippert, J.B., Holdo, R.M. (2015). Challenging the maximum rooting depth paradigm in grasslands and savannas. *Functional Ecology*, 29, 739–745. https://doi.org/10.1111/1365-2435.12390

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, 444–456. https://doi.org/10.1016/j.agrformet.2017.10.009

O

Otkin, J.A., Anderson, M.C., Hain, C., Mladenova, I.E., Basara, J.B., Svoboda, M. (2013). Examining rapid onset drought development using the thermal infrared-based evaporative stress index. *Journal of Hydrometeorology,* 14 (4), 1057–1074. https://doi.org/10.1175/JHM-D-12-0144.1.

Otkin, J.A., Anderson, M.C., Hain, C., Svoboda, M. (2014). Examining the relationship between drought development and rapid changes in the evaporative stress index. *Journal of Hydrometeorology,* 15 (3), 938–956. https://doi.org/10.1175/JHM-D-13-0110.1.

Otkin, J.A., Shafer, M., Svoboda, M., Wardlow, B., Anderson, M.C., Hain, C., Basara, J. (2015). Facilitating the use of drought early warning information through interactions with agricultural stakeholders. *Bulletin of the American Meteorological Society*, 96 (7), 1073–1078. https://doi.org/10.1175/BAMS-D-14-00219.1.

Otkin, J.A., Anderson, M.C., Hain, C., Svoboda, M., Johnson, D., Mueller, R., Tadesse, T., Wardlow, B., Brown, J. (2016). Assessing the evolution of soil moisture and vegetation conditions during the 2012 United States flash drought. *Agricultural and Forest Meteorology,* 218-219, 230–242. http://www.sciencedirect.com/science/article/pii/S0168192315300265.

Olson, M.E., Soriano, D., Rosell, J.A., Anfodillo, T., Donoghue, M.J., Edwards, E.J., León-

Gómez, C., Dawson, T., Camarero Martínez, J.J., Castorena, M., Echeverría, A., Espinosa, C.I., Fajardo, A., Gazol, A., Isnard, S., Lima, R.S., Marcati, C.R., Méndez-Alonzo, R. (2018). Plant height and hydraulic vulnerability to drought and cold. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 115, 7551–7556. https://doi.org/10.1073/pnas.1721728115

P

Pendergrass, A.G., Meehl, G.A., Pulwarty, R., Hobbins, M., Hoell, A., AghaKouchak, A., Bonfils, C.J.W., Gallant, A.J.E., Hoerling, M., Hoffmann, D., Kaatz, L., Lehner, F., Llewellyn, D., Mote, P., Neale, R.B., Overpeck, J.T., Sheffield, A., Stahl, K., Svoboda, M., Wheeler, M.C., Wood, A.W., Woodhouse, C.A. (2020). Flash droughts present a new challenge for subseasonal-to-seasonal prediction. *Nature Climate Change,* 10, 191–199. https://doi.org/10.1038/s41558-020-0709-0

Potter, S., Harrison, S., Kreft, P. (2021). The Benefits and Challenges of Implementing Impact-Based Severe Weather Warning Systems: Perspectives of Weather, Flood, and Emergency Management Personnel. *Weather, Climate, and Society,* 13, 303–314. https://doi.org/10.1175/WCAS-D-20-0110.1

Power, S., Casey, T., Folland, C., Colman, A., & Mehta, V. (1999). Inter-decadal modulation of the impact of ENSO on Australia. *Climate Dynamics*, 15, 319–324. https://doi.org/10.1007/s003820050284

Pozzi, W., Sheffield, J., Stefanski, R., Cripe, D., Pulwarty, R., Vogt, J.V., Heim, R.R., 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, 776–785. https://doi.org/10.1175/BAMS-D-11-00176.1

Prudhomme, C., Barker, L.J., Cammalleri, C., Harrigan, S., Ionita, M., Vogt, J. (2024). Chapter 13 - Drought Early Warning Systems: monitoring and forecasting, in: Tallaksen, L.M., van Lanen, H.A.J. (Eds.), Hydrological Drought (Second Edition). *Elsevier,* pp. 595–635. https://doi.org/10.1016/B978-0-12-819082-1.00002-3

Pulwarty, R. S., & Sivakumar, M. V. K. (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

Q

Qiu, J., Crow, W.T., Nearing, G.S., Mo, X., Liu, S. (2014). The impact of vertical measurement depth on the information content of soil moisture times series data. *Geophysical Research Letters*, 41, 4997–5004. https://doi.org/10.1002/2014GL060017

R

Rammig, A., Wiedermann, M., Donges, J.F., Babst, F., von Bloh, W., Frank, D., Thonicke, K., Mahecha, M.D. (2015). Coincidences of climate extremes and anomalous vegetation responses: comparing tree ring patterns to simulated productivity. *Biogeosciences*, 12, 373–385. https://doi.org/10.5194/bg-12-373-2015

Rajsekhar, D., Singh, V. P., & Mishra, A. K. (2015). Multivariate drought index: An information theory-based approach for integrated drought assessment. *Journal of Hydrology,* 526, 164–182. https://doi.org/10.1016/j.jhydrol.2014.11.031

Renzullo, L.J., van Dijk, A.I.J.M., Perraud, J.-M., Collins, D., Henderson, B., Jin, H., Smith, A.B., McJannet, D.L. (2014). Continental satellite soil moisture data assimilation improves root-zone moisture analysis for water resources assessment. *Journal of Hydrology,* 519, 2747–2762. https://doi.org/10.1016/j.jhydrol.2014.08.008

Retallack, A., Finlayson, G., Ostendorf, B., Clarke, K., & Lewis, M. (2023). Remote sensing for monitoring rangeland condition: current status and development of methods. *Environmental and Sustainability Indicators*, *19*, 100285. https://doi.org/10.1016/j.indic.2023.100285

Rickert, K.G., Stuth, J.W. and McKeon, G.M. (2000). Modelling pasture and animal production In 'Field and Laboratory Methods for Grassland and Animal Production Research'. (Eds. L.T. Mannetje and R.M. Jones), pp. 29-66 (CABI publishing: New York)

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, 381–394. https://doi.org/10.1175/BAMS-85-3-381

Rouse, J.W., 1974. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation (No. NASA-CR-139243). https://ntrs.nasa.gov/citations/19740022555

S

Sandholt, I., Rasmussen, K., & Andersen, J. (2002). A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. *Remote Sensing of environment*, *79*(2-3), 213-224. https://doi.org/10.1016/S0034-4257(01)00274-7

[Dataset] Schaaf, Crystal, Wang, Zhuosen (2015). MCD43A4 MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF Adjusted Ref Daily L3 Global - 500m V006. https://doi.org/10.5067/MODIS/MCD43A4.006

Schaefer, G.L., Cosh, M.H., Jackson, T.J. (2007). The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN). *Journal of Atmospheric and Oceanic Technology,* 24, 2073–2077. https://doi.org/10.1175/2007JTECHA930.1

Schwalm, C.R., Anderegg, W.R.L., Michalak, A.M., Fisher, J.B., Biondi, F., Koch, G., Litvak, M., Ogle, K., Shaw, J.D., Wolf, A., Huntzinger, D.N., Schaefer, K., Cook, R., Wei, Y., Fang, Y., Hayes, D., Huang, M., Jain, A., Tian, H. (2017). Global patterns of drought recovery. *Nature*, 548, 202–205. https://doi.org/10.1038/nature23021

Sheffield, J., Wood, E.F. (2012). Drought: Past Problems and Future Scenarios. Routledge. https://doi.org/10.4324/9781849775250.

Shyrokaya, A., Pappenberger, F., Pechlivanidis, I., Messori, G., Khatami, S., Mazzoleni, M., & Di Baldassarre, G. (2024). Advances and gaps in the science and practice of impact-based forecasting of droughts. *Wiley Interdisciplinary Reviews: Water*, *11*(2), e1698. https://doi.org/10.1002/wat2.1698

Skelton, R.P., West, A.G., Dawson, T.E. (2015). Predicting plant vulnerability to drought in biodiverse regions using functional traits. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 112, 5744–5749.

https://doi.org/10.1073/pnas.1503376112

Sloat, L.L., Gerber, J.S., Samberg, L.H. *et al.* Increasing importance of precipitation variability on global livestock grazing lands. *Nature Climate Change,* 8, 214–218 (2018). https://doi.org/10.1038/s41558-018-0081-5

Smith, W. K., Reed, S. C., Cleveland, C. C., Ballantyne, A. P., Anderegg, W. R. L., Wieder, W. R., ... & Running, S. W. (2019). Large divergence of satellite and Earth system model estimates of global terrestrial CO_2 fertilization. *Nature Climate Change*, 9(7), 566–571. https://doi.org/10.1038/s41558-019-0406-x

Stagge, J.H., Kohn, I., Tallaksen, L.M., Stahl, K. (2015). Modeling drought impact occurrence based on meteorological drought indices in Europe. *Journal of Hydrology,* 530, 37–50. https://doi.org/10.1016/j.jhydrol.2015.09.039

Sun, Y., Frankenberg, C., Wood, J.D., Schimel, D.S., Jung, M., Guanter, L., Drewry, D.T., Verma, M., Porcar-Castell, A., Griffis, T.J., Gu, L., Magney, T.S., Köhler, P., Evans, B., Yuen, K. (2017). OCO-2 advances photosynthesis observation from space via solar-induced chlorophyll fluorescence. *Science,* 358, eaam5747. https://doi.org/10.1126/science.aam5747

Sun, J., Zhang, B., Pan, Q., Liu, W., Wang, X., Huang, J., ... & Han, X. (2023). Non-linear response of productivity to precipitation extremes in the Inner Mongolia grassland. *Functional Ecology,* 37(6), 1663-1673. https://doi.org/10.1111/1365-2435.14328

Sutanto, S.J., van der Weert, M., Blauhut, V., Van Lanen, H.A.J. (2020). Skill of large-scale seasonal drought impact forecasts. *Natural Hazards and Earth System Sciences,* 20, 1595–1608. https://doi.org/10.5194/nhess-20-1595-2020

Sutanto, S.J., Weert, M. van der, Wanders, N., Blauhut, V., Lanen, H.A.J.V. (2019). Moving from drought hazard to impact forecasts. *Nature Communications,* 10, 1–7. https://doi.org/10.1038/s41467-019-12840-z

Swain, S., Wardlow, B.D., Narumalani, S., Rundquist, D.C., Hayes, M.J. (2013). Relationships between vegetation indices and root zone soil moisture under maize and soybean canopies in the US Corn Belt: a comparative study using a close-range sensing approach. *International Journal of Remote Sensing,* 34, 2814–2828. https://doi.org/10.1080/01431161.2012.750020

T

Tadesse, T., Demisse, G.B., Zaitchik, B., Dinku, T. (2014). Satellite-based hybrid drought monitoring tool for prediction of vegetation condition in Eastern Africa: A case study for Ethiopia. *Water Resources Research*, 50, 2176–2190. https://doi.org/10.1002/2013WR014281

Tadesse, T., Wardlow, B.D., Hayes, M.J., Svoboda, M.D., Brown, J.F. (2010). The Vegetation Outlook (VegOut): A New Method for Predicting Vegetation Seasonal Greenness. *GIScience & Remote Sensing*, 47, 25–52. https://doi.org/10.2747/1548-1603.47.1.25

Tadesse, T., Demisse, G.B., Zaitchik, B., Dinku, T. (2014). Satellite-based hybrid drought monitoring tool for prediction of vegetation condition in Eastern Africa: A case study for Ethiopia. *Water Resources Research,* 50, 2176–2190. https://doi.org/10.1002/2013WR014281

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, S., Dijk, A.I.J.M.V., Tregoning, P., Renzullo, L.J. (2019). Forecasting dryland vegetation condition months in advance through satellite data assimilation. *Nature Communications,* 10, 469. https://doi.org/10.1038/s41467-019-08403-x

Tian, S., Renzullo, L.J., Pipunic, R.C., Lerat, J., Sharples, W., Donnelly, C. (2021). Satellite soil moisture data assimilation for improved operational continental water balance prediction. *Hydrology and Earth System Sciences*, 25, 4567–4584. https://doi.org/10.5194/hess-25-4567-2021

Tootoonchi, F., Sadegh, M., Haerter, J. O., Räty, O., Grabs, T., & Teutschbein, C. (2022). Copulas for hydroclimatic analysis: A practice‐oriented overview. *Wiley Interdisciplinary Reviews: Water*, *9*(2), e1579. https://doi.org/10.1002/wat2.1579

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment,* 8, 127–150. https://doi.org/10.1016/0034-4257(79)90013-0

U

Ukkola, A. M., De Kauwe, M. G., Pitman, A. J., Best, M. J., Abramowitz, G., & Decker, M. (2016). Land surface models systematically overestimate the intensity, duration and magnitude of seasonal-scale evaporative droughts. *Environmental Research Letters,* 11(10), 104012. https://doi.org/10.1088/1748-9326/11/10/104012

V

van den Dool, H., Huang, J., Fan, Y. (2003). Performance and analysis of the constructed analogue method applied to U.S. soil moisture over 1981–2001. *Journal of Geophysical Research: Atmospheres,* 108. https://doi.org/10.1029/2002JD003114

van Hateren TC, Chini M, Matgen P, Teuling AJ. Ambiguous Agricultural Drought: Characterising Soil Moisture and Vegetation Droughts in Europe from Earth Observation. *Remote Sensing,* 2021; 13(10):1990. https://doi.org/10.3390/rs13101990

Van Dijk, A.I.J.M. (2010). The Australian Water Resources Assessment System. Technical Report 3. Landscape Model (version 0.5) Technical Description. CSIRO: Water for a Healthy Country National Research Flagship. https://awo.bom.gov.au/assets/notes/publications/Van_Dijk_AWRA05_TechReport3.pdf

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, 1040–1057. https://doi.org/10.1002/wrcr.20123

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., Beguería, S., Lorenzo-Lacruz, J., Camarero, J.J., López-Moreno, J.I., Azorin-Molina, C., Revuelto, J., Morán-Tejeda, E., Sanchez-Lorenzo, A. (2012). Performance of Drought Indices for Ecological, Agricultural, and Hydrological Applications. *Earth Interactions,* 16, 1–27. https://doi.org/10.1175/2012ei000434.1

Vicente-Serrano, S.M., Gouveia, C., Camarero, J.J., Begueria, S., Trigo, R., Lopez-Moreno, J.I., Azorin-Molina, C., Pasho, E., Lorenzo-Lacruz, J., Revuelto, J., Moran-Tejeda, E., Sanchez-Lorenzo, A. (2013). Response of vegetation to drought time-scales across global land biomes. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 110, 52–57. https://doi.org/10.1073/pnas.1207068110

Viney, Neil; Vaze, Jai; Crosbie, Russell; Wang, Bill; Dawes, Warrick, (2015). AWRA-L v5.0: Technical description of model algorithms and inputs. *CSIRO*: EP154705. http://hdl.handle.net/102.100.100/136446

Vogel, E., Lerat, J., Pipunic, R., Frost, A.J., Donnelly, C., Griffiths, M., Hudson, D., Loh, S. (2021). Seasonal ensemble forecasts for soil moisture, evapotranspiration and runoff across Australia. *Journal of Hydrology*, 601, 126620. https://doi.org/10.1016/j.jhydrol.2021.126620

W

Wang, X., Biederman, J.A., Knowles, J.F., Scott, R.L., Turner, A.J., Dannenberg, M.P., Köhler, P., Frankenberg, C., Litvak, M.E., Flerchinger, G.N., Law, B.E., Kwon, H., Reed, S.C., Parton, W.J., Barron-Gafford, G.A., Smith, W.K. (2022). Satellite solar-induced chlorophyll fluorescence and near-infrared reflectance capture complementary aspects of dryland vegetation productivity dynamics. *Remote Sensing of Environment,* 270, 112858. https://doi.org/10.1016/j.rse.2021.112858

Wang, X., Xie, H., Guan, H., Zhou, X. (2007). Different responses of MODIS-derived NDVI to root-zone soil moisture in semi-arid and humid regions. *Journal of Hydrology,* 340, 12–24. https://doi.org/10.1016/j.jhydrol.2007.03.022

Wang, Y., Zhou, H., Huang, J., Yu, J., Yuan, Y. (2023). A framework for identifying propagation from meteorological to ecological drought events. *Journal of Hydrology,* 625, 130142. https://doi.org/10.1016/j.jhydrol.2023.130142

Ward, R., Lackstrom, K., Davis, C. (2021). Demystifying Drought: Strategies to Enhance the Communication of a Complex Hazard. *Bulletin of the American Meteorological Society*, 1, 1–43. https://doi.org/10.1175/BAMS-D-21-0089.1

[Dataset] Water and Atmospheric Resources Monitoring Program. Illinois Climate Network. (2020). Illinois State Water Survey, 2204 Griffith Drive, Champaign, IL 61820-7495. http://dx.doi.org/10.13012/J8MW2F2Q.

Wedd, R., Alves, O., de Burgh-Day, C., Down, C., Griffiths, M., Hendon, H. H., ... & Zhou, X. (2022). ACCESS-S2: the upgraded Bureau of Meteorology multi-week to seasonal prediction system. *Journal of Southern Hemisphere Earth Systems Science,* 72(3), 218-242. https://www.publish.csiro.au/es/ES22026

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

Wilhite, D. A., & Pulwarty, R. S. (2005). Drought as hazard: understanding the natural and social context. In D. Wilhite (Ed.), Drought and Water Crises: Science, Technology, and Management Issues (pp. 3–29). CRC Press. https://doi.org/10.1007/S11269-006-9076-5

Wilhite, D. A., Sivakumar, M. V., & Pulwarty, R. (2014). Managing drought risk in a changing climate: The role of national drought policy. Weather and climate extremes, 3, 4-13. https://doi.org/10.1016/j.wace.2014.01.002

Wilson, L., Bende-Michl, U., Sharples, W., Vogel, E., Peter, J., Srikanthan, S., ... & Bellhouse, J. (2022). A national hydrological projections service for Australia. *Climate Services*, *28*, 100331. https://doi.org/10.1016/j.cliser.2022.100331

WMO (2015). WMO guidelines on multi-hazard impact-based forecast and warning services. WMO Doc. 1150, 34 pp., https://library.wmo.int/doc_num.php?explnum_id=7901.

WMO & GWP. (2016). Handbook of Drought Indicators and Indices (Integrated Drought Management Programme, Integrated Drought Management Tools and Guidelines Series No. 2). World Meteorological Organization and Global Water Partnership.

World Livestock: Transforming the livestock sector through the Sustainable Development Goals (2019). FAO. https://doi.org/10.4060/ca1201en

X

Xiang, K., Wang, B., Li Liu, D., Chen, C., Waters, C., Huete, A., & Yu, Q. (2023). Probabilistic assessment of drought impacts on wheat yield in south-eastern Australia. *Agricultural Water Management*, *284*, 108359. https://doi.org/10.1016/j.agwat.2023.108359

Xu, Y., Zhang, X., Hao, Z., Singh, V.P., Hao, F. (2021). Characterization of agricultural drought propagation over China based on bivariate probabilistic quantification. *Journal of Hydrology,* 598, 126194. https://doi.org/10.1016/j.jhydrol.2021.126194

Xue, J., Su, B. (2017). Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *Journal of Sensors,* 2017, 1–17. https://doi.org/10.1155/2017/1353691

Y

Yang, Y., Donohue, R. J., & McVicar, T. R. (2016). Global estimation of effective plant rooting depth: Implications for hydrological modeling. *Water Resources Research*, *52*(10), 8260-8276. https://doi.org/10.1002/2016WR019392

Yilmaz, M. T., Hunt Jr, E. R., & Jackson, T. J. (2008). Remote sensing of vegetation water content from equivalent water thickness using satellite imagery. *Remote Sensing of Environment*, *112*(5), 2514-2522. https://doi.org/10.1016/j.rse.2007.11.014

Yuan, X., Wood, E.F. (2013). Multimodel seasonal forecasting of global drought onset. *Geophysical Research Letters,* 40, 4900–4905. https://doi.org/10.1002/grl.50949

Yuan, W., Zheng, Y., Piao, S., Ciais, P., Lombardozzi, D., Wang, Y., ... & Myneni, R. (2019). Increased atmospheric vapor pressure deficit reduces global vegetation growth. *Science Advances,* 5(8), eaax1396. https://doi.org/10.1126/sciadv.aax1396

Yuan, X. (2016). An experimental seasonal hydrological forecasting system over the Yellow River basin–Part 2: The added value from climate forecast models. *Hydrology and Earth System Sciences,* 20(6), 2453-2466. https://doi.org/10.5194/hess-20-2453-2016

Z

Zambrano, F., Vrieling, A., Nelson, A., Meroni, M., Tadesse, T. (2018). Prediction of drought-induced reduction of agricultural productivity in Chile from MODIS, rainfall estimates, and climate oscillation indices. *Remote Sensing of Environment,* 219, 15–30. https://doi.org/10.1016/j.rse.2018.10.006

Zhang, Y., Zhou, S., Gentine, P., & Xiao, X. (2019). Can vegetation optical depth reflect changes in leaf water potential during soil moisture dry-down events?. *Remote Sensing of Environment*, *234*, 111451. https://doi.org/10.1016/j.rse.2019.111451

Zeng, L., Wardlow, B.D., Xiang, D., Hu, S., Li, D. (2020). A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sensing of Environment,* 237, 111511. https://doi.org/10.1016/j.rse.2019.111511

Zeng, J., Zhou, T., Qu, Y., Bento, V. A., Qi, J., Xu, Y., ... & Wang, Q. (2023). An improved global vegetation health index dataset in detecting vegetation drought. Scientific Data, 10(1), 338. https://doi.org/10.1038/s41597-023-02255-3

Zhang, Y., Zhang, Y., Lian, X., Zheng, Z., Zhao, G., Zhang, T., ... & Piao, S. (2023). Enhanced dominance of soil moisture stress on vegetation growth in Eurasian drylands. *National Science Review*, 10(8), nwad108. https://doi.org/10.1093/nsr/nwad108

Zscheischler, J., Westra, S., van den Hurk, B. J. J. M., Seneviratne, S. I., Ward, P. J., Pitman, A., ... & Leonard, M. (2018). Future climate risk from compound events. *Nature Climate Change*, 8(6), 469–477. https://doi.org/10.1038/s41558-018-0156-3