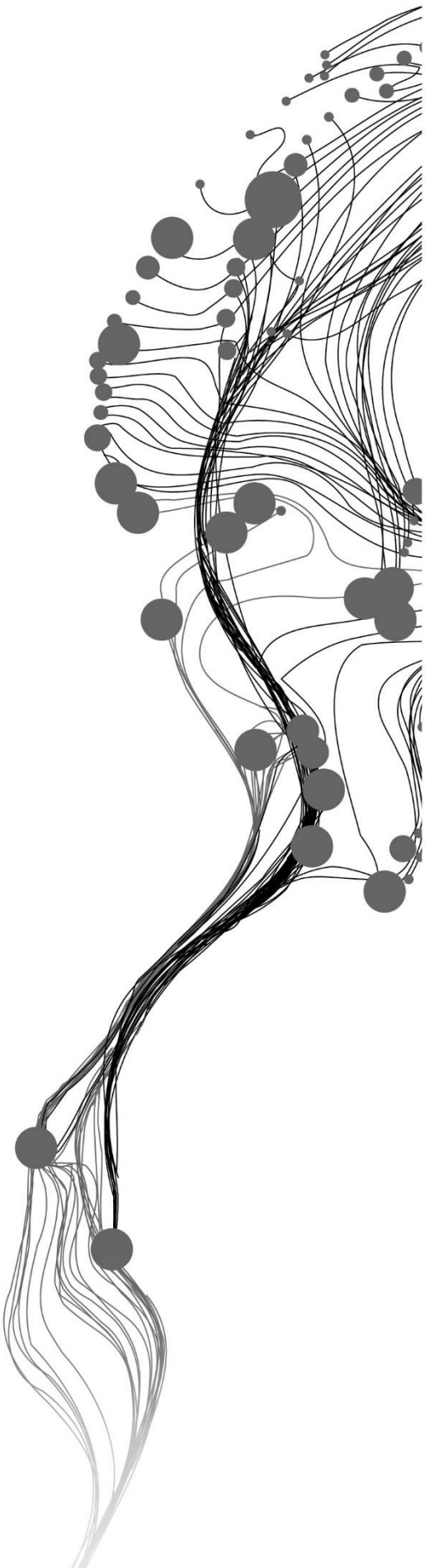
UNRAVELING THE SPATIAL-TEMPORAL DYNAMICS OF DROUGHT ANOMALIES AND THEIR INTERACTIONS WITH VEGETATION.

CALVIN SAMWEL SWAI
July 2024

SUPERVISORS:
Dr, Y, Zeng
Dr.ir, S, Salama



CALVIN SAMWEL SWAI
Enschede, The Netherlands, July 2024

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Water Resources and Environmental Management

SUPERVISORS:
Dr, Y, Zeng
Dr.ir, S, Salama

THESIS ASSESSMENT BOARD:
Dr, M.T, Marshall (Chair)
Prof, S, Manfreda (External Examiner, University of Napels)

UNRAVELING THE SPATIAL-TEMPORAL DYNAMICS OF DROUGHT ANOMALIES AND THEIR INTERACTIONS WITH VEGETATION.

ABSTRACT

The future projections sketched in the various Climate reports predict a global increase in severe weather conditions, particularly excessive decline both in summer precipitation and winter snowfall accompanied by periods of extremely high temperatures. Consequently, these projections are expected to intensify the frequency and intensity of drought events affecting thereby water and food security on a global scale. Due to expected local shifts in weather conditions, understanding the effect on drought dynamics at a local catchment will be critical for coordinated risk management and informed mitigations of drought impacts on both water resources and the ecosystem. However, complexities in the development of drought scenarios limit the capacity to establish a conclusive link of drought propagation in a local area. Moreover, the uneven distribution of the Root Zone Soil Moisture (RZSM) dataset also limits the assessment of the relationship between drought and vegetation anomalies.

The purpose of this study is to analyze the spatial-temporal dynamics of drought anomalies and their causal relationships with vegetation conditions in Overijessel, Netherlands. To address the issue of limited coverage of the RZSM dataset, the readily available data on Surface Soil Moisture at 1km (1KmSSM) was validated with the in-situ Twente soil moisture measurements. These data were then used to calibrate the two models to derive RZSM from Surface Soil Moisture (SM). The Cumulative Distribution Function (CDF) and Exponential Low Pass Filter (ELPF) models. The optimal calibrated model was used to retrieve the RZSM time series in the Witharen groundwater monitoring network (location: Twente the Netherlands). Thereafter, the estimated RZSM time series, along with precipitation measurements, and groundwater time series were used to estimate soil moisture and meteorological, and groundwater anomalies for drought monitoring in the region. Additionally, optical vegetation indices Near-Infrared reflectance of vegetation (NIRv) and Normalized Difference Water Index (NDWI) were used to monitor both the anomalies of the vegetation photosynthetic activities and hydrology dynamics during various drought conditions. Finally, a Convergent Cross-Mapping (CCM) was used to assess the causal relationships between drought and vegetation anomalies in the Witharen region.

The findings of this study showed that the ELPF method estimated the most optimal RZSM time series from 1KmSSM in Twente stations, with a mean unbiased Root Mean Square Error (ubRMSE) of 0.056 m3/m3. The temporal analysis of drought anomalies indicated a higher drought frequency and intensity in the last decade, while spatial analysis revealed an association between groundwater depth and the development of groundwater drought. Issues such as delayed response and periodic recovery of vegetation anomalies during drought events were identified in the time series analysis. Moreover, the detected significant drought interactions aligned with the prior knowledge of the Witharen region. For instance, the interaction between shallow groundwater levels (0.7 to 3 m) and both vegetation and soil moisture anomalies.

Moreover, this study also highlighted the use of both surface and root zone soil moisture anomalies for monitoring meteorological drought propagation and the initiation of vegetation and groundwater anomalies respectively. Nevertheless, this study suffered from key sources of uncertainties related to the spatial and temporal resolution of the analysis, and the accuracy of the grid soil moisture datasets. Future studies should address these issues to improve the comprehensive assessment of drought development in local areas. This study concludes that CCM can be used to reveal valuable information about the local causal interactions between drought and vegetation anomalies through time series analysis. The use of both surface and root zone soil moisture in drought analysis opens a potential to uncover significant drought interactions both on the surface and sub-surface.

Keywords

Drought, Root Zone Soil Moisture, Convergent Cross-Mapping, Spatial-Temporal Dynamics.

ACKNOWLEDGEMENTS

I would like to extend my genuine gratitude to my supervisors – Dr. Yijian Zeng and Dr. Ir, Suhyb Salama for their priceless support and guidance throughout the Thesis writing process. They provided me with words of encouragement, mentored me, gave me constructive feedback and most importantly were patient with me as I learned. I can safely say I have learned a lot from this experience because of them. I would also like to extend my gratitude to Mostafa Daoud for helping with various questions regarding groundwater dynamics and interactions. I would also like to extend my gratitude to Dr. Eng. Rogier van der Velde and Drs. Tom Hoogland at Vitens for the help during the groundwater data collection process.

And finally, I want to extend my heartfelt thanks to my family, friends, Aquila, and my partner. Your support has been crucial, to say the least. I honestly believe I would not be where I am now here if I didn't have your support. Thank you so much.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Background

Drought is one of the most destructive natural disasters, characterized by limited water availability below the climatological average. Drought significantly impacts vegetation health and food productivity among other dimensions. Recently, a significant decline of 30% in Gross Plant Productivity (GPP) was reported in Europe during the 2003 and 2018 drought episodes (Buitink et al., 2020). Global climate projections predict a warming climate that could potentially increase drought frequency, severity, and vegetation sensitivity to drought (AghaKouchak et al., 2014; Christian et al., 2023). Additionally, the impacts on global food security are expected to be compounded by the global population increase (UN, 2022). Therefore, understanding drought causations and the dynamics with vegetation conditions is relevant for both effective water resources management and global food security.

Impacts of drought can be manifested in a range of aspects from hydrometeorological conditions and socio-economic factors that influence both water availability and its demand. Because of this, various approaches can be used for the interpretation of drought events (Rutger Weijers, 2020). The most popular approach considers the disciplinary perspective of drought impacts, to classify the four critical stages of drought severity. Wilhite et al. (1985) illustrated these stages as meteorological, agricultural, hydrological, and finally socio-economic drought. Soil moisture anomaly was identified as an indicator of agricultural drought indicating a link with vegetation health. However, Van Hateren et al. (2021) indicated the ambiguity concerning the use of the term agricultural drought. Hereafter, the aforementioned stages will be referred to as drought types and soil moisture drought will be used in place of agricultural drought for consistency.

Despite a vague separating boundary between the identified drought types, the classification provides a way to physically identify drought phenomena, quantifying their impacts and characterizing their development. Meteorological drought refers to a measure of precipitation deficiency, which when prolonged can develop into soil moisture and hydrological (also groundwater) drought as a result of critical values of soil moisture, streamflow, and groundwater (Z. Han et al., 2021; Park et al., 2021). Socio-economic drought, on the other hand, occurs in extreme drought situations leading to socioeconomic losses such as loss of lives, conflicts, and migration (Su et al., 2017).

While it is common to conceptualize the propagation of drought across the identified stages, it is far less convenient to conclude with certainty their causal interactions. Since the introduction of the drought stages, studies have used various approaches such as probabilistic linkages, time series analysis, and temporal correlation to highlight the linkages between meteorological, soil moisture, and groundwater anomalies (Dai et al., 2022; Q. Li et al., 2019; Z. Xu et al., 2023; H. Zhang et al., 2021). However, evidence in these studies highlights the temporal variation in the nature and magnitude of the relationship between drought stages complicating the drought causation conclusion.

The interplay of issues such as delayed response, accumulated impacts, and a third-factor influence has been explained to constrain drought causation conclusions. For instance, severe hydrological drought in the Netherlands during the summer of 2019 is associated with precipitation anomalies from the summer of

2018 to the early spring of 2019 (Philip et al., 2020). In another study, using a non-linear Spearman's correlation analysis H. Zhang et al. (2021) found a delayed response of soil moisture and groundwater anomalies to meteorological drought which is shorter during summer and autumn compared to spring and winter. Furthermore, Rutger Weijers (2020) explained the seasonal decoupling and changes in causation direction between precipitation, soil moisture, and groundwater dynamics.

Adding vegetation degradation to the equation further complicates the dynamics of the drought relationship. Although, drought events have been mostly associated with vegetation degradation (Su et al., 2017; Wilhite et al., 1985), their relationship has been explained to be equally complicated. A study by Buitink et al. (2020) found a significant vegetation-delayed response to critical soil moisture levels highlighting the impact of vegetation self-coping mechanisms on the nature of vegetation response (W. Zhang et al., 2023). External factors such as temperature and land elevation were also found to dictate the reaction of vegetation to drought conditions (Van Hateren et al., 2021). Additionally, other studies found that the impact of drought anomalies on vegetation is both spatially and temporally variable (AghaKouchak, Farahmand, et al., 2015; AghaKouchak, Feldman, et al., 2015).

The discussed issues have motivated a shift from linear correlation analysis into non-linear approaches for studying drought causations. The most recently adopted non-linear approach is the use of the Convergent Cross-Mapping (CCM) method. Originally proposed by Sugihara et al., (2012) for the analysis of ecological interactions between species and environmental factors, CCM has gained popularity in the field of hydrology due to its ability to detect causal relationships between hydrological variables (Bonotto et al., 2022; Kadir et al., 2020; Y. Wang et al., 2018). In drought studies, CCM has demonstrated a capacity to unravel drought causation dynamics by inferring the time series of drought anomalies which are also complex and non-linear (Ionita et al., 2022; Shi et al., 2022; Shiau, 2023).

Another notable development in drought analysis has been the application of Root Zone Soil Moisture (RZSM) to infer soil moisture drought rather than Surface Soil Moisture (SSM). Although SSM has reported advantages in monitoring soil water and energy dynamics, it lacks reliability when it comes to measuring soil water availability for plants. This is because primarily, SSM reflects the conditions on shallow soil typically about 5cm deep as explained by Cai et al., 2022 and Wakigari & Leconte, 2022. RZSM, on the other hand, is a more reliable measure of a plant's soil water availability as the saturation of the soil root zone reflects the actual amount of water available for plant roots (Grillakis et al., 2021; Zhuang et al., 2020). RZSM has also been explained to be a key indicator of plants' water stress and vegetation productivity, highlighting its applicability in vegetation drought monitoring (Carranza et al., 2021).

From the literature reviewed above, it is apparent that significant progress has been made in the analytical approaches for characterizing, quantifying, and analyzing the interaction between drought and vegetation anomalies. The promising results demonstrated by the application of CCM causal analysis to decipher time series non-linear causality interactions offer an opportunity to address non-linear links between drought types in various spatial scales. Moreover, a more physically explainable RZSM for reflecting the plant's water availability over SSM improves the prospect of understanding the link between drought and vegetation anomalies.

In the Netherlands however, the application of these developments for the monitoring of drought development is still limited. This is a concern considering the projected warming climate that aligns with the global climate projections. The KNMI'14, a 4-scenarios translation of the IPCC 2013 climate change

scenarios specific to the Netherlands projected a decline in summer precipitation and winter snowfall coupled with a temperature rise, solar radiation reaching the earth's surface, and of course the potential evapotranspiration (Van Den Hurk et al., 2014). Reports also predict a decrease in the return period of severe drought episodes similar to the 2018 drought (Philip et al., 2020). Understanding drought causation dynamics in the Netherlands is crucial, especially considering the impacts of the 2018 drought on the ecosystem conditions and water resources (Brakkee et al., 2021, 2022), to inform the development of rational mitigation measures.

The limited spatial-temporal coverage of the in situ RZSM dataset, as reported by Carranza et al. (2021) and Pezij et al. (2019), along with the available low-resolution grid dataset for instance Martens et al. (2017) has led to its underutilization in localized soil moisture drought studies. However, the development of promising less data-demanding methods to derive RZSM from the high-resolution grid SSM time series provides an opportunity for retrieving RZSM for localized drought studies (X. Gao et al., 2017; Zhuang et al., 2020). Moreover, the documented potential of using remote sensing for soil moisture and vegetation condition monitoring provides a viable solution to address the demand for spatial-temporal continuous monitoring (B. C. Gao, 1996; Q. Han et al., 2023; Zeng et al., 2022).

1.2. Research Objectives and Questions

Based on the identified research gaps, there seems to be an opportunity to improve the state of research in the Netherlands. This study was designed to address the holistic understanding of the causal dynamics of drought and vegetation anomalies in the selected area of the Netherlands. The main aim of this study is to comprehensively assess the interactions of water-induced stress with vegetation conditions. Specifically, this objective was broken down into the following sub-objectives and questions:

I. To assess the most effective method for estimation of high-resolution Root Zone Soil Moisture (RZSM) in Twente spanning 2000-2020.
    - What is the feasibility of using a 1 Km global surface soil moisture dataset (1KmSSM) for monitoring soil moisture content in the Twente region?
    - How do various methods, both analytical and statistical, capture and depict variations in RZSM trends in the Twente region?
    - Which methodology is the most suitable method for estimating RZSM spanning 2000-2020 from the 1KmSSM dataset in the study area?
    - Can the calibrated 1km RZSM estimation model be transferred and generalized with acceptable accuracy?

II. To assess the spatial-temporal dynamics of soil moisture, meteorological, and groundwater drought indices in the study area spanning the period from 2000-2020.
    - What are the spatial-temporal dynamics and trends of soil moisture, meteorological, and groundwater drought indices, in the study area?

III. To assess vegetation conditions across the study area under different water availability conditions.
    - How is the vegetation condition associated (or varies) with various water availability indices in the study area?

- What insights will be gained regarding vegetation conditions by utilizing the plant hydrology and biophysical vegetation optical indices?

IV. To unravel the cause-effect relationship between climate variability (net precipitation), water availability (drought indices), and vegetation conditions (spectral indices).

- How do water availability (groundwater, soil moisture) and the anomalies of climate variables (net precipitation) causally relate to the fluctuations of plant conditions?

This study enhances the existing knowledge of drought-based studies in the Netherlands through the following contributions:

I. Investigating the robustness of the methods for Root Zone Soil Moisture (RZSM) estimation from grid Surface Soil Moisture (SSM), enhancing the analysis of the plants' water availability.
II. Determining spatial factors for drought development and propagation, improving insights into the localized drought resilience and vulnerability for dedicated mitigation and adaptation measures.
III. Detecting causal links enhances the ability to predict and mitigate the impacts of drought on vegetation, agriculture, and water resources.

This study sets itself from the existing literature, for instance, studies by Shiau (2023) and Shi et al. (2022) by adopting a comprehensive approach to evaluate the causalities between three drought types and vegetation anomalies. Furthermore, it enhances the existing study on Root Zone Soil Moisture (RZSM) estimation and application in soil moisture drought analysis. The findings of this study offer insight into the feasibility of 1KmSSM for surface soil moisture monitoring at a local scale and RZSM estimations. Additionally, it contributes to the investigation of approaches for the RZSM estimation from satellite soil moisture datasets and addresses spatial-temporal issues associated with drought development complexities. Finally, through the application of a non-linear model, this study aims to contribute to drought propagation analysis by uncovering non-linear causation interactions between drought types and vegetation anomalies.

2. STUDY AREA AND DATASETS

2.1. Study Area

This study was conducted in the Overijssel province of the Netherlands, focusing on two specific locations: Twente and Witharen region. In the Twente region moisture monitoring network was used for the validation of the 1KmSSM dataset as well as the calibration of the RZSM estimation method. Witharen groundwater monitoring network was used as the primary location for the analysis of drought and vegetation anomalies interactions for its dense network of groundwater monitoring wells.

Figure 1A shows the location of the Overijssel province within the Netherlands. The distribution of the soil moisture and groundwater monitoring networks together with the study area topography are shown in Figure 1 B. Soil moisture monitoring sites are placed between 5km to 13km apart and cover a rough area of 1800 km². Groundwater monitoring wells, on the other hand, are more nucleated within a rough box area of 69.4 km² separated by a distance of about 120 m to 3 km between the closest monitoring stations.

Figure 1 C and 1 D presents the land cover properties of the two monitoring locations. The main land cover properties identified in the study area include croplands, built-ups, and grasslands. Relatively, the Withren region is dominated by an extensive cropland area together with woodlands and grasslands whereas the Twente monitoring stations are mostly in built-up areas and cultivated land. Surface water bodies range from small features such as ponds, and drainage canals to rivers such as Dinkel, and Ommerkanaal river streams form the surface water system in the study area.

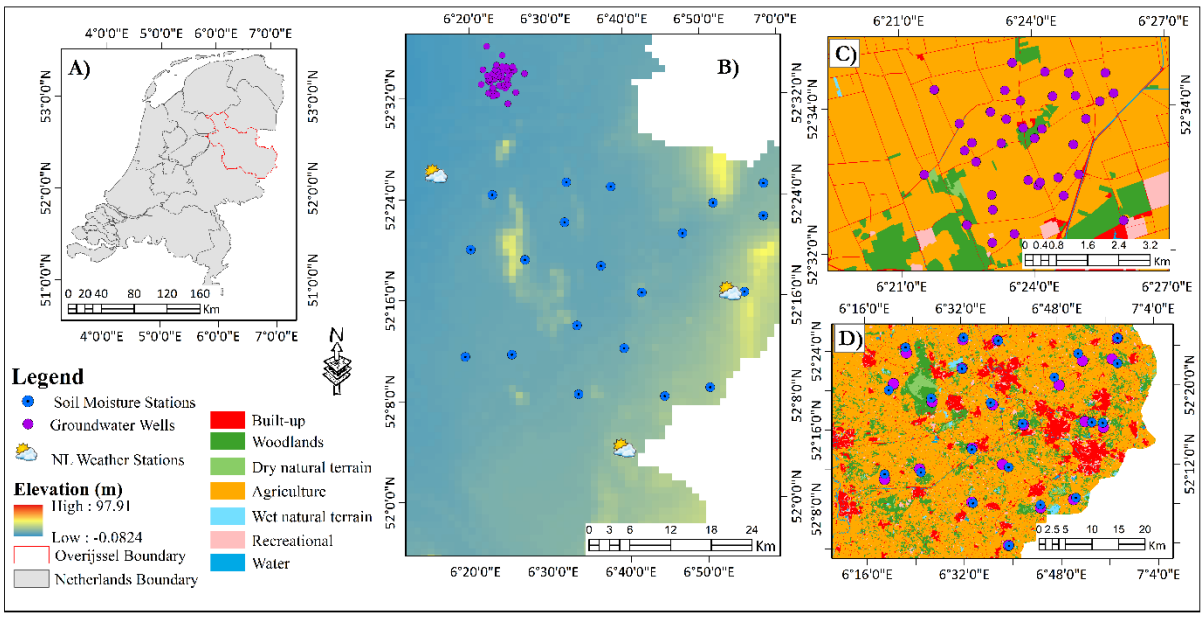


Figure 1: Study area map a) Location in Netherlands map, B) Elevation, C) and D) Land Cover Map for Witharen and Twente region

The climate of the study area is mostly influenced by the Ocean-atmospheric interactions in the west and landmass on the east side bringing moisture and heat (Van Der Velde et al., 2021). As a result, the area experiences uniformly distributed precipitation throughout the year with April (33.5mm) and August

(86.8mm) being the driest and wettest months of the year respectively. The area has an annual water surplus of approximately 145.8mm after balancing precipitation with evaporative water demand. The temperature, on the other hand, ranges between below -6^0 c to above 30^0 c in winter and summer respectively. Literature reports extreme conditions during 2018, 2019, and 2020 as the most extreme dry years experienced whilst August 2010, 2011, and 2013 were the wettest periods (Guardamino et al., 2021; Rutger Weijers, 2020; Van Den Hurk et al., 2014).

2.2. Data Sources

2.2.1. Remote Sensing datasets

In Table 1, an overview of the remote sensing datasets utilized in this study is provided. Classified as remote sensing products include primary satellite products such as MODIS MOD09GA dataset as well as secondary spatial continuous datasets such as Bofek 2020 Soil Map and 1KmSSM.

Table 1: An overview of the remote Sensing datasets utilized in this study.

Dataset	Used Variable	Rationale	Source, Access	Spatial Resolution	Temporal Resolution
Global Surface Soil Moisture (1KmSSM)	Surface Soil Moisture	Soil moisture drought monitoring	(Q. Han et al., 2023), Earth Engine Catalog	1 Km	Daily, 2000-2020.
ECMWF Reanalysis v5 (ERA5)	Precipitation and Actual Evapotranspiration Dataset	Meteorological drought index computation	European Centre for Medium-Range Weather Forecasts (ECMWF), Earth Engine Catalog	11132 m	Daily, 1979-Now.
Netherlands Land Cover Maps	Land cover classes	The analysis of land cover properties.	Statistics Nederland and Top10NL	1:10000	Not Applicable.
Soil Map (Bofek 2020)	Soil texture and hydrological properties	Analysis of soil properties	(Schröder et al., 2021), Netherlands Data Portal	1:50000	Not Applicable.
Actuell Hoogetebestand Nederland (AHN3)	Topographical information; Elevation and slope	Topography analysis	(Actuell Hoogetebestand Nederland, 2015), Earth Engine Catalog	0.5 m	Not Applicable.
MOD09GA	Land surface reflectance.	Retrieving Vegetation indices (NDWI, NIRv)	United States Geological Survey (USGS), Earth Engine Catalog	500 m	2000-Now

2.1.1 In-situ Measurements

2.1.1.1 Soil Moisture Measurements.

Soil moisture measurements in the Twente region were used as the in-situ soil moisture inputs for the calibration of RZSM estimation models and evaluation of 1KmSSM accuracy. The Twente soil moisture monitoring site includes 22 soil moisture monitoring stations installed in various locations around the Overijssel province (figure 1). Measurements are taken at five soil depths by the probes installed at the surface (5cm deep), 10 cm, 20 cm, 40 cm, and 80 cm from the ground surface. Van Der Velde et al. (2022) published a calibrated and corrected dataset that includes soil moisture and soil temperature measurements at the 5 depths for the period 2008-2020 at a fine temporal scale of 15 minutes. However, due to maintenance issues and instrumentation limitations, soil moisture measurements are not consistently available at all depths. In this study, 15 out of 22 soil moisture stations were retrieved for their consistency in dataset availability at the depths of 5 to 40 cm for an extended period of at least 5 years (2016-2020).

2.1.1.2 Meteorological variables.

In-situ measurements of precipitation were required to validate the use of a spatially continuous ERA5 dataset. In this study, precipitation measurements from the Koninklijk Nederlands Meteorologish Instituut (KNMI) were used to test the accuracy of the ERA5 precipitation dataset. KNMI is the Dutch Royal Meteorological institution operating the system of automatic weather stations in the Netherlands with data from 1951 to the present. Time series of daily validated precipitation measurements for the 3 KNMI stations close to the study area i.e., Heino, Hupsel, and Twenthe were retrieved via the KNMI database.

2.1.1.3 Groundwater Measurements.

Groundwater time series datasets were retrieved from the publicly available Dinoloket soil research database. The database was developed and managed by the TNO Geological Survey of the Netherlands is the center for accessing deep and shallow underground data and models. Groundwater datasets including geographical location, water elevation, filter depth, and surface elevation were downloaded in a CSV format. The temporal resolution of groundwater level measurements was available in 14 days while 3-hour measurements became available in 2006.

Out of 42 available groundwater monitoring wells in Witharen, 27 wells were selected to be used in this study due to their consistency and temporal coverage. Groundwater analysis focused on the time series retrieved from the phreatic filters (between 0.7 to 3 m deep from the ground surface). Figure 2 shows the distribution of the 20-year average groundwater depth for the retrieved time series. Most of the monitoring wells are shallow with a water table depth suggesting a significant interaction between groundwater and surface water dynamics should be expected.

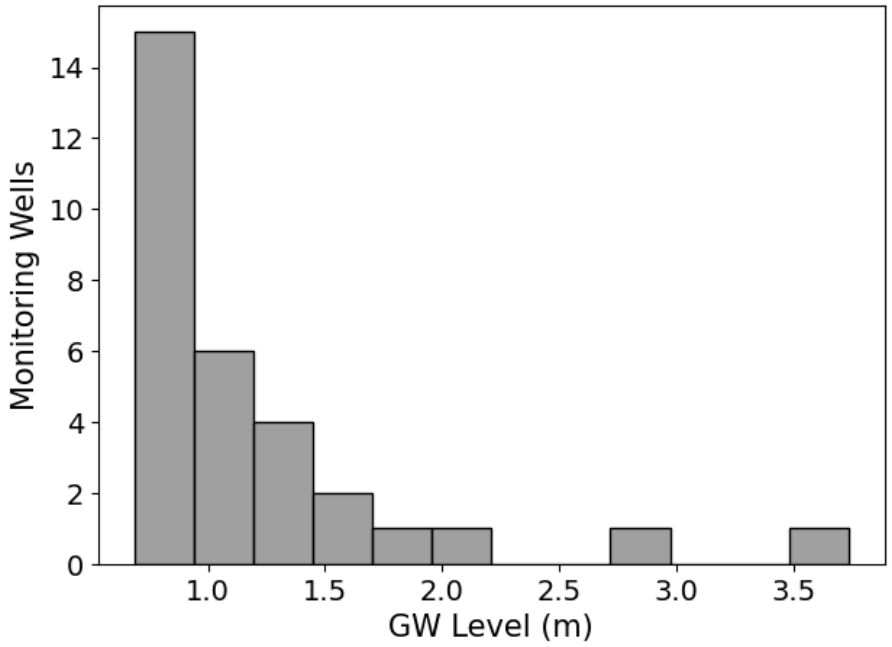


Figure 2: Histogram showing the distribution of the groundwater depth for the monitoring wells

3 DATA PROCESSING AND METHODOLOGY APPROACH

3.1 Data Processing and Preparation

This study used datasets from various sources, leading to challenges caused by the heterogeneity of data formats and resolutions in addition to individual data errors. In this section, two main data processing procedures are presented in detail. First, the processing and preparation of the optical remote sensing datasets together with the resolution matching are illustrated. Second, the accuracy of the grid datasets ERA5 and 1KmSSM time series are presented together with the process used to prepare in-situ Root Zone Soil Moisture (RZSM) time series. It should be noted that preparation procedures for each dataset including outlier treatment and unit conversions presented in their respective objectives in the methodology section.

3.1.1 Optical Data Processing

Land surface reflectance information was derived from MOD09GA version 6.1. A multispectral optical image dataset containing 7 surface reflectance spectral bands has been available since February 24, 2000. An atmospherically corrected dataset for aerosols, gases, and aerosols was derived from the Google Earth Engine code editor. In addition to the constituent bands that are designed within specific wavelength ranges to study various properties on the earth's surface. Additionally, data quality pixels are also available to provide information about pixel quality (Roger et al., 2020).

Before computing vegetation indices, MOD09GA surface reflectance was processed to remove unwanted pixels. Pixels polluted by clouds and snow cover, together with empty pixels were identified using 'state_1km' and 'num_observations_1km' bands respectively. Clouds, cloud shadow, and snow cover labels were all retrieved from the 'state_1km' band, by using specific Bits (1,2, and 10-15). Empty pixels were identified by label 0 in the 'num_observations_1km' band. A custom-made clouds-empty pixels mask was created in the Google Earth engine code editor in which unwanted clouds, cloud shadow, and snow-covered pixels together with empty pixels were removed from the final image collection to be used for vegetation indices computation.

3.1.2 Resolution Matching

As indicated in the previous section, datasets used in this study were retrieved from their respective databases, each with its own temporal and spatial properties. This sub-section illustrates the matching process as well as the rationale used to align the temporal and spatial resolution of the data inputs.

Grid format datasets i.e., 1KmSSM, MOD09GA, ERA5, and AHN3 were originally retrieved with the spatial resolution of 1Km, 500 m, 11 Km, and 0.5 m respectively. Spatial downscaling of soil moisture and precipitation datasets requires the integration of auxiliary datasets that account for the factors influencing the hydrological cycle at a target resolution (Bai et al., 2019; Chen et al., 2020). Data quality issues associated with optical datasets have been reported to influence uncertainty to the downscaled products while microwave high-resolution datasets are limited in a long-term study (Sang et al., 2022).

Consequently, this study utilized a simple median to rather upscale MOD09GA extractions into a 1 Km resolution matching 1KmSSM dataset. A spatial median of the neighboring pixels was used as it provides

an average condition of the pixels while avoiding complexities that can be introduced by regressing models. At each location of the monitoring station of interest, a buffer of 500 m was used to extract the median value for the pixel neighboring pixels. Extractions from a course resolution ERA5 dataset (11 Km) and vector format soil map could not be rescaled to 1 Km resolution using the median method. For that case, a less variable precipitation trend was assumed across the 69 Km2 areas of the Witharen monitoring wells, and soil information was assessed at a 500 m buffer scale.

On the other hand, the temporal resolution of the retrieved datasets was in 2-week and 3-hours records for groundwater measurements, 15 minutes for in-situ soil moisture, and daily for ERA5, MOD09GA, and 1KmSSM datasets. In the first objective in-situ soil moisture was aggregated into a daily mean to match daily records of 1KmSSM datasets. Clouds and cloud shadow pixel cleaning led to the loss of daily availability of the optical data MOD09GA, resulting in an average 13-day temporal resolution. Therefore, for the second to the fourth objective, drought indices computation as well as the causal analysis was conducted on a monthly resolution to ensure maximum consistency of the temporal resolution.

3.1.3 Remote Sensing Dataset Validation

1KmSSM Validation.

In this subsection, the accuracy of the 1KmSSM gridded surface soil moisture dataset is presented before application in RZSM estimations. The assessment was conducted on the 15 monitoring stations with 2016-2020 soil moisture measurements consistently available in all 4 soil depth levels. RMSE and correlation coefficients were used to assess the accuracy of the 1KmSSM in terms of linear agreement and residual measurements respectively (Equations 4 and 6).

Figure 3 presents the distribution of the two-accuracy metrics with respect to temporal-matched surface soil moisture measurements. Quantitative measures of the 1KmSSM performance indicate an agreement with in-situ soil moisture measurements in most stations. Specifically, Figure 3 B reveals a median correlation coefficient of 0.57 between 1KmSSM and in-situ surface soil moisture measurements highlighting the capacity of the 1KmSSM to match the temporal dynamics of the in-situ surface soil moisture.

However, the results of the residual measurements by RMSE suggest a significant deviation of 1KmSSM from the actual SSM measurements. The median RMSE of 0.1 m^3/m^3 provides a general overview of the average difference between 1KmSSM and in-situ soil moisture measurements. A minimum RMSE of 0.07 m^3/m^3 indicates that in all monitoring stations, 1KmSSM failed to accurately reproduce surface soil moisture incidences at a desirable measure of residuals of 0.04 m^3/m^3 (Van Der Velde et al., 2021).

Following the accuracy assessment of 1KmSSM, seven stations that scored a moderately strong correlation (>0.6) with in-situ soil moisture were used for the calibration and evaluation of RZSM estimation methods.

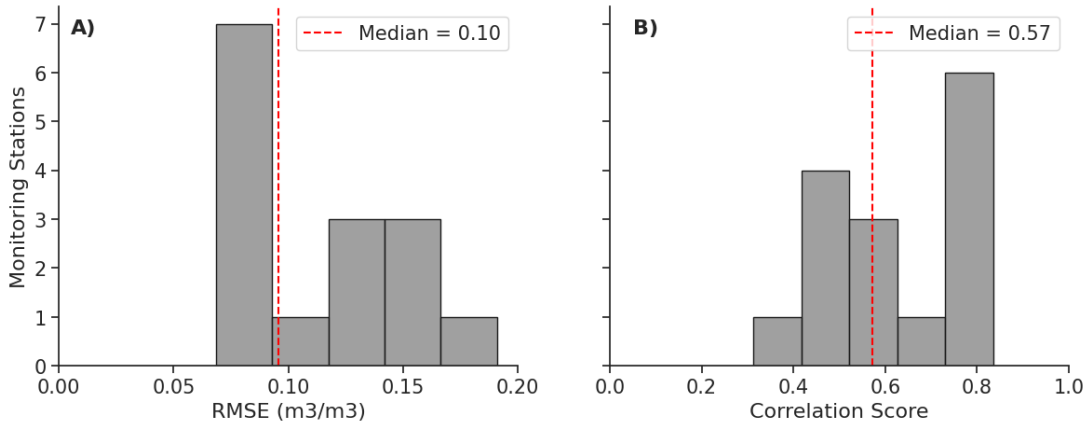


Figure 3: 1KmSSM accuracy matrices A) RMSE and B) Correlation Score

In Figure 4, the time series of the 1KmSSM and in-situ surface soil moisture for the three of the selected stations are demonstrated. Visible gaps in the soil moisture time series are from the missing in-situ soil moisture dates, labeled -99.999 on data files of the calibrated time series published by Van Der Velde et al. (2022). Evidently, 1KmSSM time series showed agreeing seasonality dynamics respective of in-situ measurements hence the correlation coefficient >0.60.

However, extreme values of in-situ surface soil moisture, including peak values and low measurements were not replicated by 1KmSSM. In stations ITC_SM02 and 07, the 1KmSSM time series consistently underestimated the surface soil moisture peaks in winter 2018 and 2020 while in dry periods of summer 2016, 2018, and 2019 1KmSSM overestimated soil moisture content for station ITC_SM13. The deviations in critical measurements account for the residual distance measure as indicated by the average RMSE>0.1 m^3/m^3.

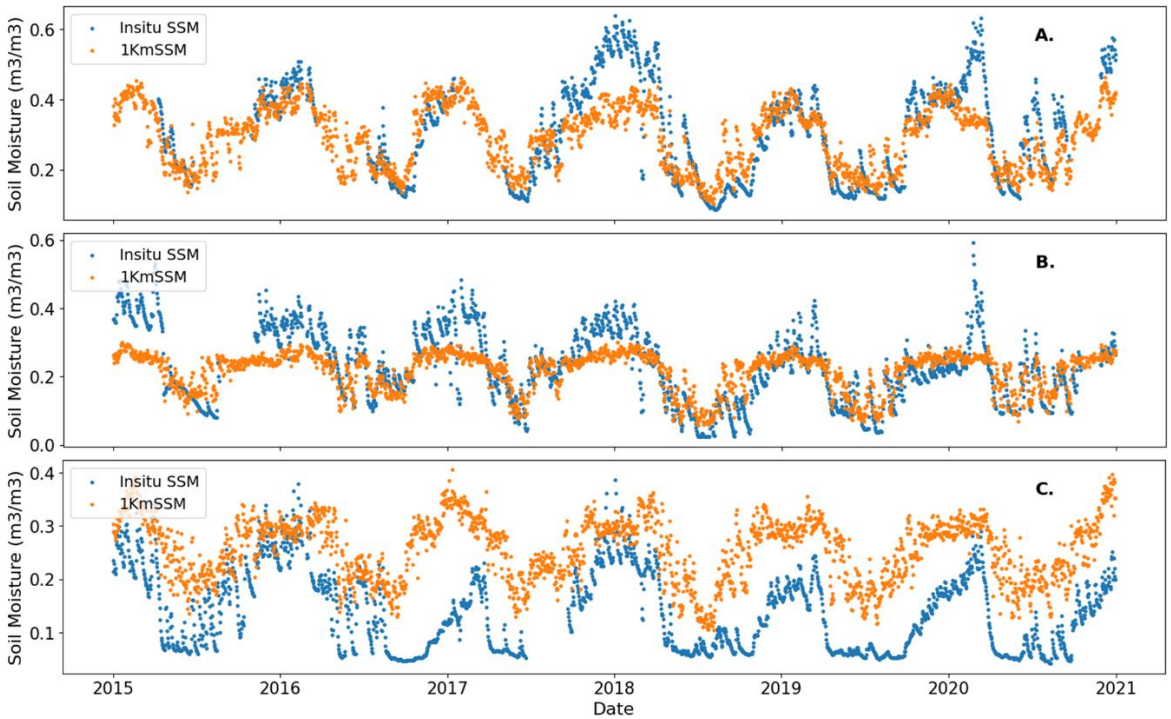


Figure 4: 1KmSSM vs In-situ surface soil moisture (5cm) time series for stations ITC_SM02, 07, and 13 (A, B, and C respectively).

ERA5 Validation

This subsection validates the accuracy of the ERA5 dataset to represent the meteorological information of the study area using in-situ measurements retrieved from KNMI stations. The validation process was conducted by quantifying the linear correlation and average deviation of ERA5 precipitation measurements (stored in band '**total_precipitation_sum'**) using correlation coefficient and Root Mean Square Error (RMSE) respectively (Equations 4 and 6). Table 2 below summarizes the scores of the two-accuracy metrics, correlation score, and RMSE used for the validation of ERA5 precipitation ('**total_precipitation_sum'**) against corresponding KNMI precipitation measurements by stations closest to the study area.

The statistical metrics from the validation analysis indicated ERA5 precipitation dataset adequately replicated the KNMI precipitation measurements. Specifically, the ERA5 precipitation dataset achieved correlation and RMSE values between 0.74 to 0.78 and 2.67 to 3.26 mm, respectively, across all three KNMI stations. The lowest correlation score of 0.74 was scored by Hupsel station which also scored 3.265mm RMSE value, the weakest performance among the three stations. Further analysis indicated that the anomaly detection improved an agreement up to 0.86 correlation coefficient between SPI-1 anomalies computed from ERA5 and the KNMI precipitation dataset over the study period (Appendix figure S1).

Table 2: Root Mean Square Error and Correlation scored by ERA5 precipitation dataset in the three KNMI weather stations between 2000-2020.

Metric	**Heino**	**Twente**	**Hupsel**
RMSE (mm)	2.68	2.8	3.26
Correlation score	0.78	0.78	0.74

3.1.4 Root Zone Soil Moisture Computation

Twente soil moisture stations were used as the in-situ measurements (available for 5cm, 10cm, 20cm, and 40cm) for the estimation of RZSM. The threshold depth of 40 cm depth was used to account for the identified depth with the significant distribution of perennial crop roots. According to Fan et al. (2016), about 70% of the roots for crops such as grains and cereals are distributed at about 30 cm depth suggesting a significant depth for RZSM monitoring for the one leading agricultural products in the Netherlands (Ellen, 1992).

RZSM in this study was defined as the mean value of soil moisture content in the depths of 0 to 40cm. However, soil moisture measurements at Twente stations are only available at the 4 selected soil depths (figure 5). For that case, to account for the unrepresented depths, a depth scaling approach in equation 1 was used to compute $RZSM_{in\text{-}situ}$ from the discrete depth measurements.

$$RZSM_{in-situ} = \frac{\sum_{j=1}^{n} \Theta_j \delta Z_j}{z} \quad \text{equation 1}$$

Where Θ_j is the volumetric water content (in m^3/m^3) measured at a depth of j (cm), z (cm) is the total depth (cm) and δZ_j is the soil thickness associated with the Θ_j measurements. At each measuring depth, the measured soil moisture was associated with half of the soil thickness from the following station.

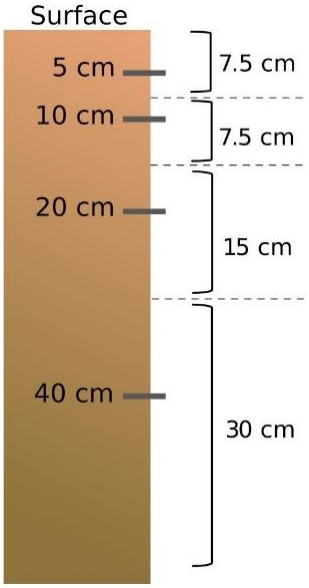


Figure 5: Demonstration of the soil moisture measurements setup across different depths. (Carranza et al., 2021)

3.2 Methodology

To assess the relationship between drought and vegetation conditions, a causal analysis was conducted using vegetation and drought anomaly time series spanning 2000-2020. In this section, a methodology design is presented that shows the stages of the conducted study divided into various objectives to achieve the causal analysis of drought anomalies. The flowchart demonstrating the overview of the methodological design is presented in Figure 6 below.

i. **Investigation of the best RZSM estimation model**: For the first objective, two RZSM estimation methods were calibrated and evaluated to determine the best-performing model using Twente monitoring stations.
ii. **Drought indices computation and analysis**: In the second objective, meteorological and groundwater drought indices were computed from meteorological and groundwater datasets respectively. Soil moisture drought was computed using the RZSM time series retrieved using the best-selected method in the first objective.
iii. **Vegetation indices and anomaly computation**: For the third objective, optical-based vegetation indices were computed from the optical remote sensing dataset.
iv. **Causal analysis**: Finally, for the fourth objective, causal interactions between the computed drought indices together with vegetation anomalies time series were analyzed.

The setting for the first objective was in Twente soil moisture monitoring stations, while a high network of groundwater monitoring stations in Witharen justified the shift in the causal relation analysis from the second to the fourth objective. A detailed explanation of the methods is explained in the sub-sections below.

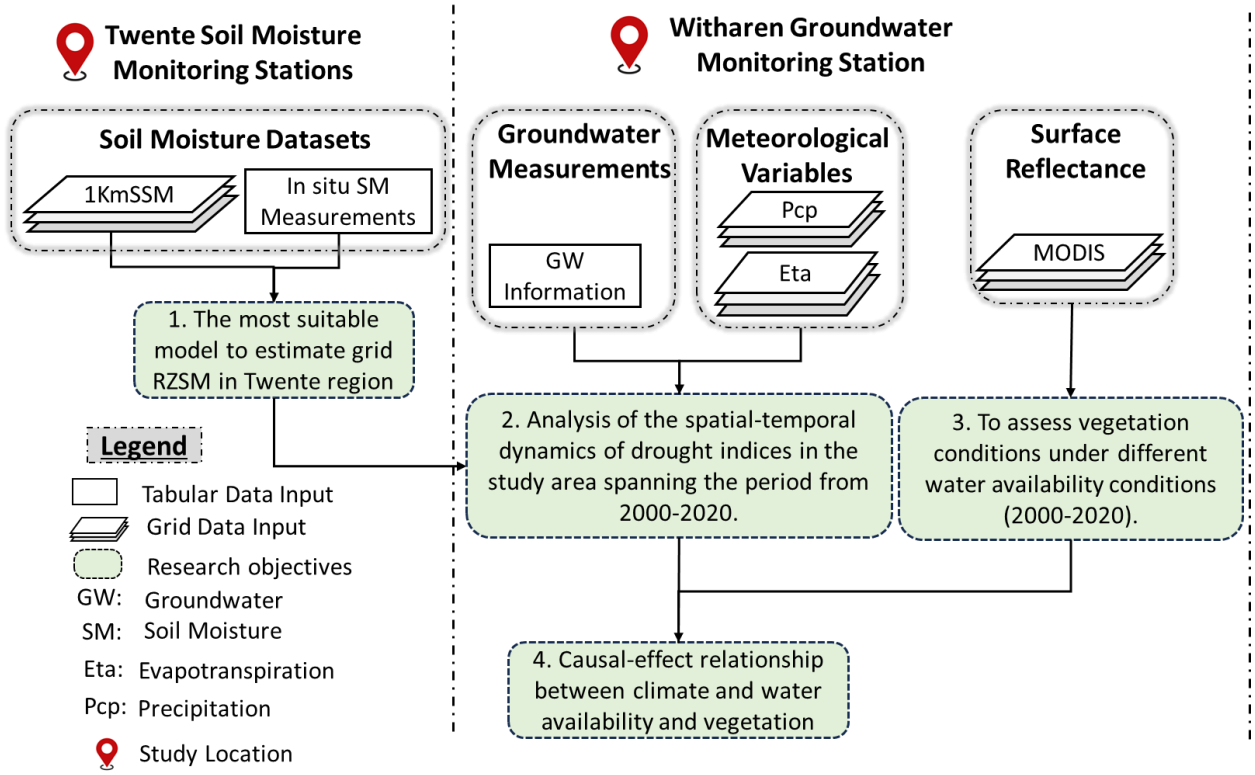


Figure 6: Demonstration of the overall methodological design adopted in this study

3.2.1 RZSM Estimation Methods

In this subsection, two methods for RZSM estimation from 1KmSSM: Exponential Low Pass Filter (ELPF) and Cumulative Distribution Function (CDF) matching are presented. Figure 7 provides a demonstration of the steps and methodology adopted to achieve this objective. Before the modeling process, the in-situ soil moisture and the extracted 1KmSSM time series per location were preprocessed for consistency. Missing values treatment and daily mean aggregation were conducted for the already pre-processed in-situ soil moisture measurements to produce a clean daily time series. For the 1KmSSM, the pixel values are stored in 1000*m3/m3 so, the actual soil moisture time series was extracted by dividing pixel values by 1000. Finally, the complete training set was created from the temporal matchups of the daily aggregates of in-situ soil moisture measurements and the extracted 1KmSSM time series.

From the complete training set, five-year time series (2016-2020) with consistent in-situ soil moisture measurements at the depths of 5 to 40 cm were prepared for the 7 selected soil moisture stations (Section 3.1.3). Time series splitting into calibration and validation sets included 2016, 2017, and 2018 data inputs and 2019 and 2020 respectively. X. Gao et al. (2017) indicated that 2-year daily measurements of soil moisture are enough to create sets for both ELPF and CDF model parameter fitting and 1-year dataset for the evaluation. A simple splitting procedure was adopted in this study to ensure each set included both severely dry years (2018 and 2019) and less dry years (2016, 2017, and 2020) (Buitink et al., 2020; Van Der Velde et al., 2022). Finally, the best method was selected between the two evaluated models as the one that produces the most desirable accuracy metrics. Method transferability was also tested for the best-performing model using two uncalibrated Twente soil moisture stations.

A detailed explanation of the working principles and steps taken for both CDF and ELPF methods is provided in the sub-sections below.

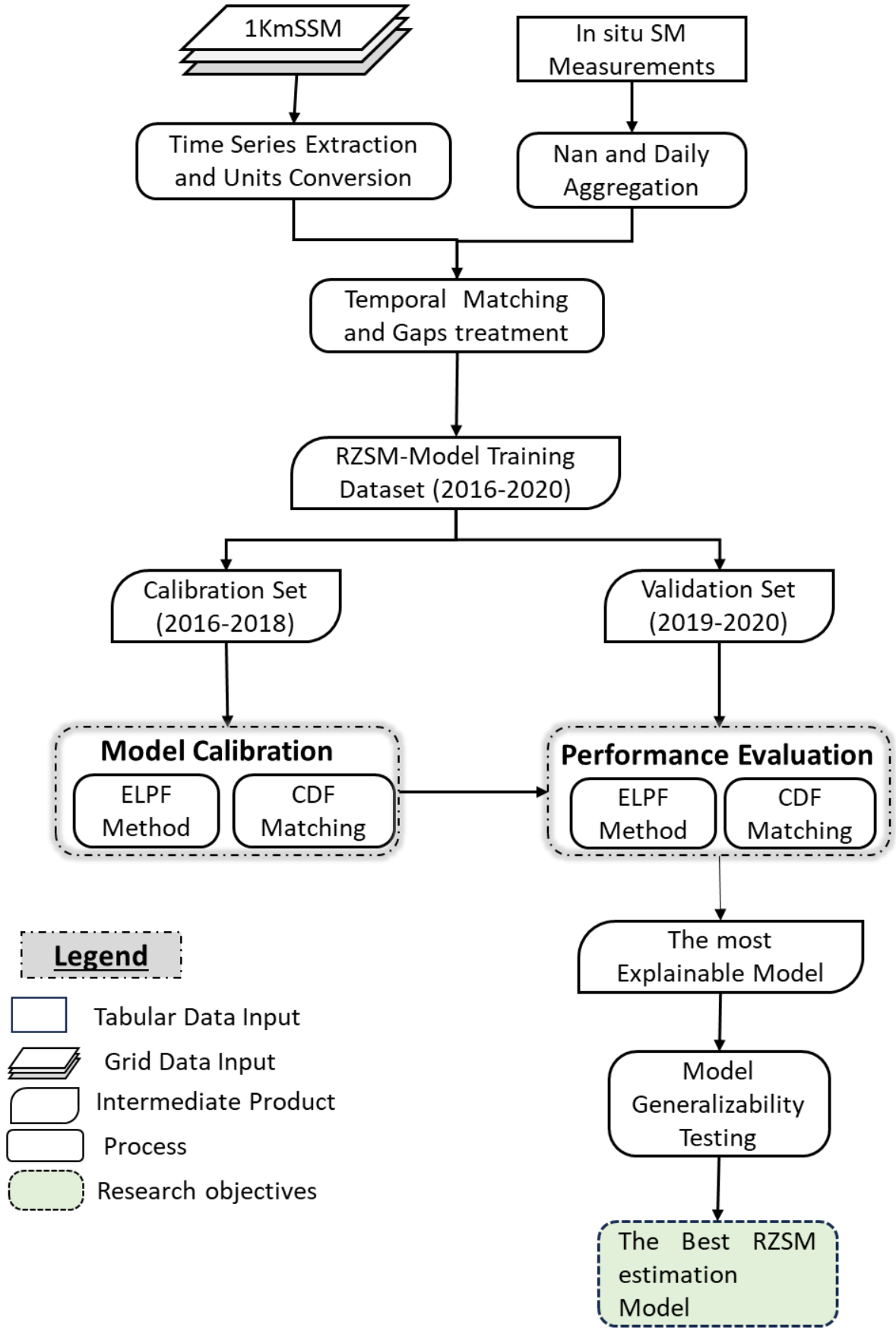


Figure 7: Demonstration of the methodological approach applied in the first objective

3.2.1.1 Cumulative Distribution Function (CDF) Matching Method

CDF matching method estimates RZSM by eliminating systematic differences between grid Surface Soil Moisture (1KmSSM) and in-situ Root Zone Soil Moisture ($RZSM_{in\text{-}situ}$) in terms of residuals (Zhuang et al., 2020). A line-fitting approach was used to explain the difference between 1KmSSM and $RZSM_{in\text{-}situ}$ dynamics using the 1KmSSM dynamics as the predicting variable. A successful CDF matching results in the matching cumulative frequencies for the predicted and the observed RZSM time series (figure 8).

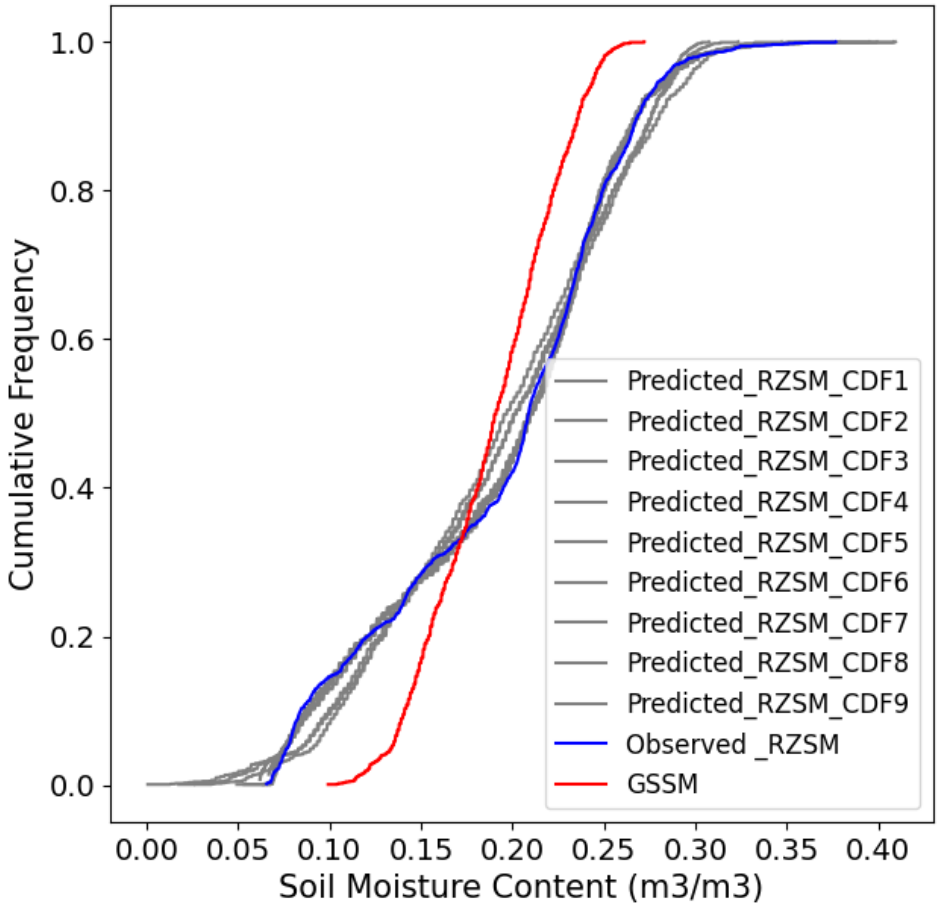


Figure 8:Demonstration of the CDF matching results.

To estimate the RZSM time series from 1KmSSM, the following steps were adopted in this study:

I. After the splitting process, both 1KmSSM and $RZSM_{in\text{-}situ}$ calibration sets were ranked in descending order. The ranking process is done to match the percentiles of the two-time series (1KmSSM and $RZSM_{in\text{-}situ}$), to reduce their systematic differences (L. Wang & Chen, 2014).

II. Ranked difference (Δ) was computed between ranked $RZSM_{in\text{-}situ}$ and 1KmSSM.

$$\Delta = 1KmSSM - RZSM_{in\text{-}situ}.$$

III. Fitting a polynomial line between ranked difference and ranked 1KmSSM on the calibration set. Various polynomial lines were calibrated to find a method that better predicts the ranked difference using the ranked 1KmSSM data inputs.

IV. Select the best-fitted polynomial line in the calibration process to estimate the Δ (Δ_{pred}) using 1KmSSM time series**.** The best-fitted polynomial line was selected as the one with the least RMSE error estimate. For instance, figure 9 demonstrates the decline in RMSE with an

increased polynomial degree, however, no significant improvement was observed after polynomial 5 for that case it was adopted as the best polynomial fit.

V. Estimate RZSM using the computed Δ_{pred} and 1KmSSM of the validation set.

$$\mathbf{RZSM_{CDF} = 1KmSSM - \Delta.}$$

VI. Steps I, IV, and V were then repeated in the validation set to estimate $RZSM_{CDF}$.

VII. Equations 4,5,6, and 7 were used to evaluate the accuracy of the $RZSM_{CDF}$ was evaluated against the $RZSM_{in\text{-}situ}$ time series in the validation dataset.

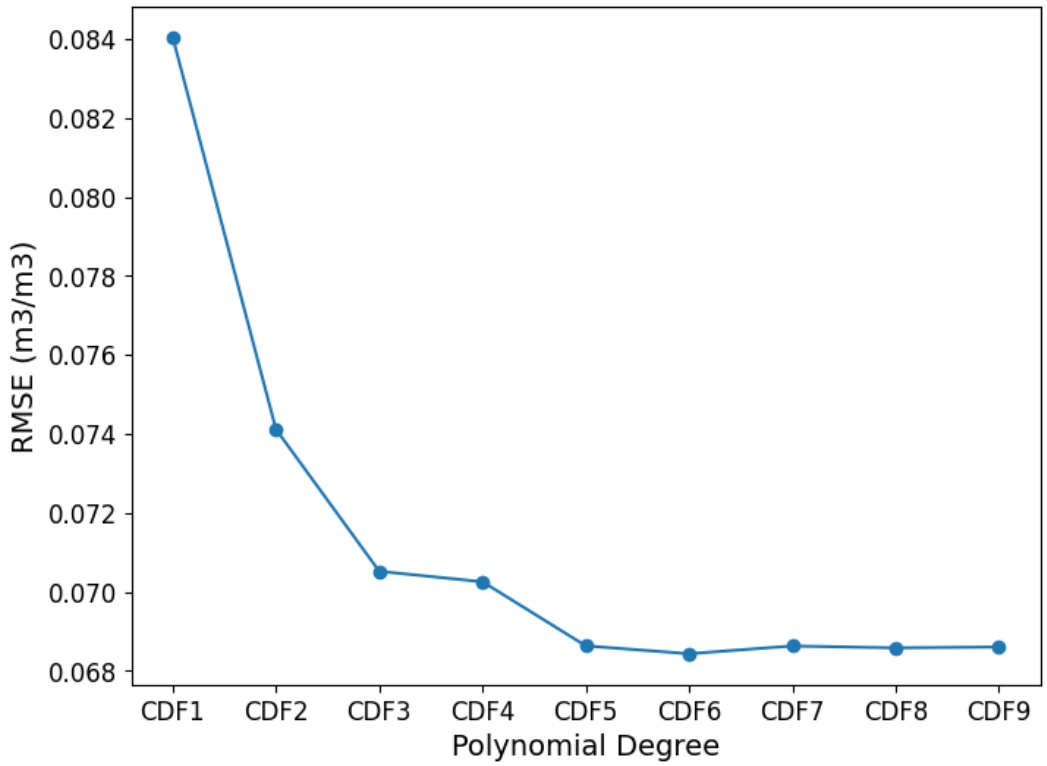


Figure 9: Demonstration of the changes in the RMSE with an increase in polynomial degree during CDF matching.

3.2.1.2 Exponential Low Pass Filter (ELPF) Smoothing Method

ELPF involves the application of the time series smoothing algorithm, delaying the upper-level soil moisture content signals to match the temporal dynamic of the deeper level. The original approach was introduced by Wagner et al., (1999) in which satellite soil moisture estimates were used to estimate profile soil moisture. In this study, however, a recursive approach introduced by (Albergel et al., 2008) as an improvement of Wagner's approach was used. In contrast to Wagner's approach, Albergel's approach uses a more simplified formulation but produces similar results to Wagner's approach.

Equation 2 below summarizes the mathematical formulation of Albergel's approach:

$$SWI_n = SWI_m + K_n * (ms_{t_n} - SWI_m) \qquad \text{equation 2}$$

Where SWI_n is the estimated deep Soil Water Index (i.e, at 10, 20, and 40 cm depth) on a day n, SWI_m is the Soil Water Index a day prior (n-1) and $ms(t_n)$ is the SSM at time n. Kn is the gain factor at time n. Kn can be expressed as equation 3 below:

$$K_n = \frac{1}{1+\sum_i^{n-1} e^{-\frac{t_n - t_m}{T}}} \qquad \text{equation 3}$$

Where $t_n and\ t_m$ are the time indices for the SWI estimation date and the day prior respectively. T, on the other hand, represents the characteristic time length taken for the surface soil moisture measurements to be translated to SWI. T parameter is also defined as the integral parameter that accounts for the soil properties and hydrological processes that influence the interaction between surface soil moisture and deep soil moisture dynamics (X. Gao et al., 2017; Wagner et al., 1999).

In the current study, the estimation of RZSM was implemented following steps

I. After the data splitting process, in the calibration phase, for a particular depth, the SWI and K_n computations were initiated by assigning SWI=SSM and Kn=1 to the first-row entry in the time series (Albergel et al., 2008).
II. SWI and K_n computations were iterated through a range of T values from 1 to 100 using equations 2 and 3 respectively and saved respectively of the T value. Paulik et al. (2014) explained that the optimal T value increases with depth, while most studies found 20 days as the optimal value for root zone depth up to 60 cm (M. Li et al., 2023).
III. The estimated SWI that correlates the most with the in-situ soil moisture measurements at the respective depth was saved as the predicted SWI and its T value was selected as the optimal T parameter for the respective depth and location.
IV. Steps I to IV were then repeated to estimate SWI at the depth of 10, 20, and 40 cm (SWI_{10}, SWI_{20}, and SWI_{40} respectively).
V. Finally, SWI_{10}, SWI_{20}, and SWI_{40} were used to estimate $RZSM_{ELPF}$ using the depth-weighted mean method introduced in Equation 1.
VI. Steps II to V were repeated for the validation set.
VII. Finally, the performance evaluation was done on the $RZSM_{in\text{-}situ}$ time series in the validation dataset using equations 4, 5, 6, and 7.

Notably, two different techniques were applied by the ELPF and CDF approaches to RZSM estimations. While CDF was calibrated and validated directly to the $RZSM_{in\text{-}situ}$, ELPF was applied to each depth independently before being depth-scaled into RZSM. The choice of the technique is based on the performance of each approach respective of the two techniques used. (Appendix figures S7 and S8 compare the performance of the two techniques in both CDF and ELPF)

3.1.1.1 RZSM Accuracy Evaluation Metrics

The accuracy of each method was evaluated to investigate the best-performing method for estimating RZSM from the existing 1KmSSM dataset. Four widely used accuracy metrics in soil moisture studies, Root Mean Square Error (RMSE), Bias term, correlation coefficient (r), and the unbiased Root Mean Square Error (ubRMSE) were utilized for the evaluation process. H. Zhang et al. (2021) utilized similar indices to provide a mathematical quantification of the average difference between the predicted and actual measurements, systematic errors in the modeling results, the strength of the linear correlation, and the average difference

in the absence of systematic errors respectively. Equations 4, 5, 6, and 7 were adopted for the computation of the respective metrics.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{n}} \quad \text{Equation 4}$$

$$Bias = \frac{\sum_{i=1}^{n}(Y_i - X_i)}{\sum_{i=1}^{n} X_i} \quad \text{Equation 5}$$

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2} * \sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \quad \text{Equation 6}$$

$$unRSME = \frac{1}{n}\sqrt{n\sum_{i=1}^{n}(Y_i - X_i)^2 - (\sum_{i=1}^{n}(Y_i - X_i)^2)} \quad \text{Equation 7}$$

3.2.2 Drought Indices Computation and Characterization:

In this objective, a time series of 3 types of drought indices, meteorological drought, soil moisture, and groundwater drought were computed. SPEI-1, SSWI-1, and SGWI-1 time series were extracted at a groundwater location to monitor the evolution of the three drought types in the Witharen groundwater monitoring region. Figure 10 presents an overview of the methodological approach as well as steps undertaken to compute these drought indices.

Before drought anomalies computations, data processing steps were applied to 1KmSSM, groundwater, precipitation, and evapotranspiration separately.

- Groundwater time series underwent data exploration procedures using time series plotting and histogram plots to detect outliers. Extreme groundwater values were replaced by the 2.5th and 97.5th percentile scores of the groundwater level for the respective stations. (Demonstration: Appendix figure S2)
- Time series extraction and unit conversion used in the first objective were adopted to prepare the 1KmSSM time series before applying the RZSM estimation model calibrated in the first objective to retrieve the RZSM time series.

Finally, together with the extracted precipitation and evapotranspiration time series, both datasets were monthly aggregated using the monthly median score of the respective variables before temporally matching to generate a complete drought analysis dataset which was fed into drought computation models explained below.

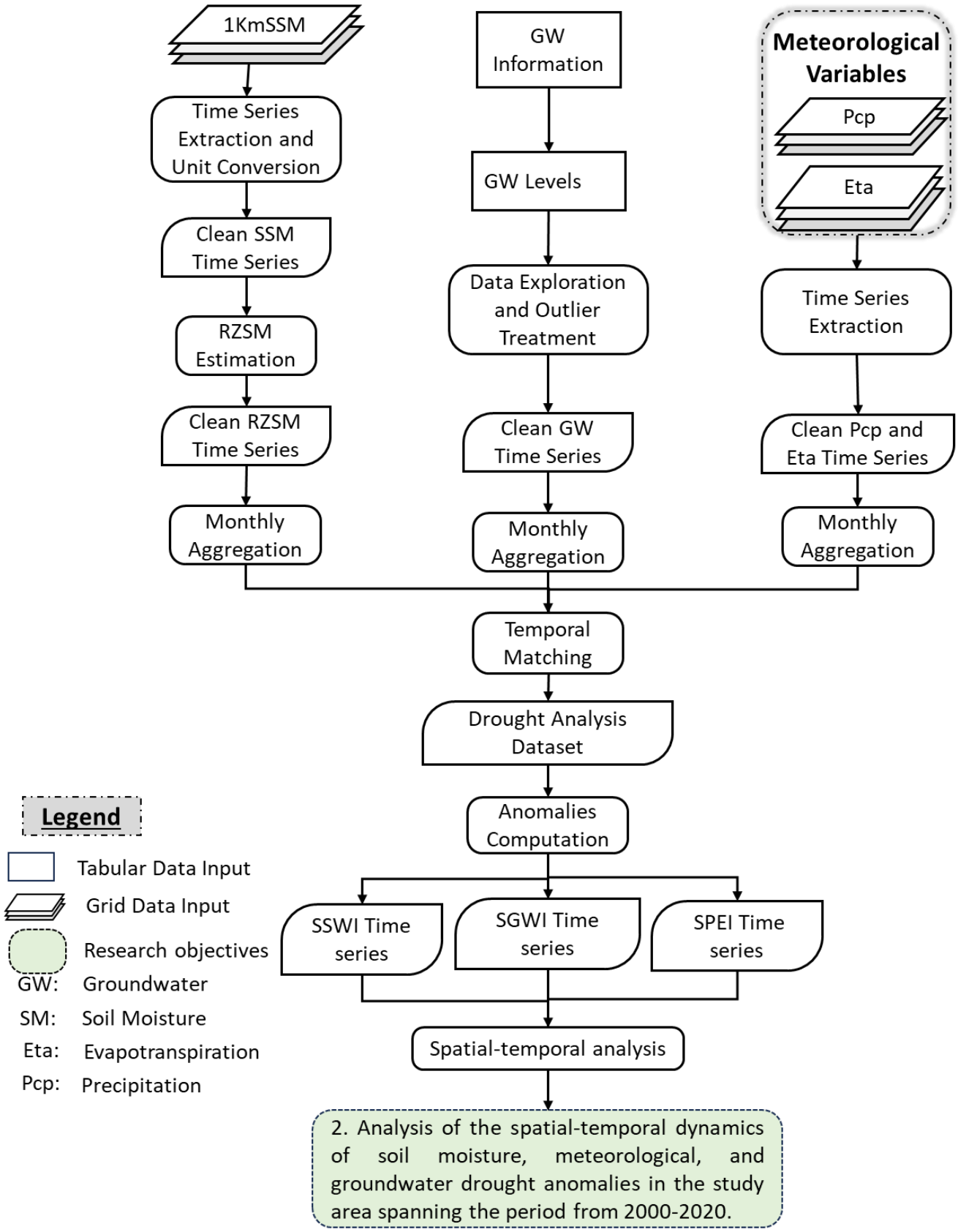


Figure 10: Demonstration of the methodological approach applied in the second objective

Computation and interpretation of drought indices are based on various temporal scales, most commonly 1, 3, 6, 9, 12, or 24 months (Su et al., 2017). Guardamino et al. (2021) indicated an increase in the correlation of the meteorological drought index with soil moisture and groundwater drought indices as one increases the monthly scale for the computation of SPEI. This is because a precipitation deficit over an extended period will impact the soil water storage. Rutger Weijers (2020) explained the use of SPEI-3 for meteorological drought propagation to soil moisture and groundwater drought respectively. In addition to the SPEI-1, SSWI-1, and SGWI-1 for the monitoring of the onset and duration of drought events, the current study also utilized SPEI-3 to account for the accumulated precipitation deficit in the causation analysis (Section 4.4.1).

Below a detailed explanation of the approaches used for each drought index is presented.

3.1.1.2 Standardized Precipitation Evapotranspiration Index (SPEI) Computation

SPEI is an energy-based atmospheric water demand index for the estimation of meteorological drought. The computation of SPEI is based on the difference between the recorded precipitation intensity and the unit Evapotranspiration measure also referred to as net precipitation. The steps for the computation of SPEI followed these of SPI with one key difference of using the net precipitation over the actual precipitation (Guardamino et al., 2021). In the current study, the following main steps were taken after data collection and preprocessing:

I. A moving average computation on the selected time scale of 3 months.
II. Fitting the climatic records of net precipitation and normalizing the skewed distribution to a normal distribution with mean and standard deviation of 0 and 1 respectively.
III. Computing the anomalies from the normalized distribution using the general standardization equation below.

$$\frac{(x_i - x_o)}{\sigma_{(x)}} \qquad \text{equation 5}$$

In which x_i is the newly computed average over the identified time scale, x_o is the mean score for the whole time series and $\sigma_{(x)}$ is the standard deviation of the selected time scale compared to the rest of the time series.

3.2.2.1 Standardized Soil Water Index (SSWI) Computation

SSWI was used as an index for soil moisture drought monitoring. RZSM time series extracted using the 1KmSSM surface soil moisture time series was used as an input for the estimation of SSWI. The standardized equation 6 below was used to compute SSWI in the study period spanning 2000-2020:

$$SSWI = \frac{RZSM_i - RZSM_\alpha}{RZSM_\sigma} \qquad \text{equation 6}$$

Whereby SSWI is the Standardized Soil Water Index, RZSMi is the RZSM at a given time step, RZSMα is the mean value in the time step and RZSMσ is the standard deviation for the RZSM value in the time step.

Additionally, SSM anomaly (SSWI-1u) was computed using 1KmSSM time series instead of RZSM. SSWI-1u was used in the causal analysis (Section 4.4.2) to assess the ability of surface soil moisture anomaly to capture drought propagation in comparison with RZSM anomalies

3.2.2.2 Standard Groundwater Level Index (SGWI) Computation

The Standardized Groundwater Level Index (SGWI) was computed to detect the onset of groundwater level declines as well as the temporal trend (Halder et al., 2020). In this study, SGWI was computed on a temporal scale and classified as indicated by Table 3. SGWI will be estimated using equation 7 below:

$$SGWI = \frac{GWL_i - GWL_\propto}{GWL_\sigma} \qquad \text{equation 7}$$

SGWI is the standardized Groundwater Level Index, GWLi is the groundwater level at a given time step, GWLα is the mean value in the time step and GWLσ is the standard deviation for groundwater level in the time step.

3.2.2.3 Drought Characterization

Unlike other hydrological phenomena such as floods, it is often less straightforward to precisely identify properties such as the onset and offset of drought events (Wanders et al., 2010). In this study, run theory was used to quantitatively classify characteristic features of drought events using the time series of the three computed drought indices (W. Zhang et al., 2023). Initially, meteorological, soil moisture, and groundwater drought events (DE) were identified based on the magnitude of the net precipitation, soil moisture, and groundwater departure from the climatological mean respectively (Su et al., 2017). For this case, a 20-year climatological average. Table 3 presents the drought classifications and corresponding standard deviation measures as identified by WMO (2012).

Table 3: Drought classification according to SPI values

DEVIATION RANGE	ASSIGNED VALUE	DROUGHT CATEGORY
F ≥ μ+2σ	≥ 2.00	Extremely wet
μ+1σ to μ+1.99σ	1 to 1.99	Wet
μ-0.99σ to μ+0.99σ	-0.99 to 0.99	Near Normal
μ-0.99σ to μ-1.99σ	-1 to -1.99	Dry
F ≥ μ-2σ	≤2.00	Extremely dry

Thereafter, drought duration (DD) was identified as the period from the occurrence to the termination of the classified dry anomaly, whereas Drought Severity (DS) was computed as the absolute sum of the respective drought index and drought intensity (DI) as the ratio of DD and DS. Moreover, Drought Frequency (DF) was calculated as the number of drought events over the time scale of interest. The demonstration of the drought characterization approach is presented in Figure 11.

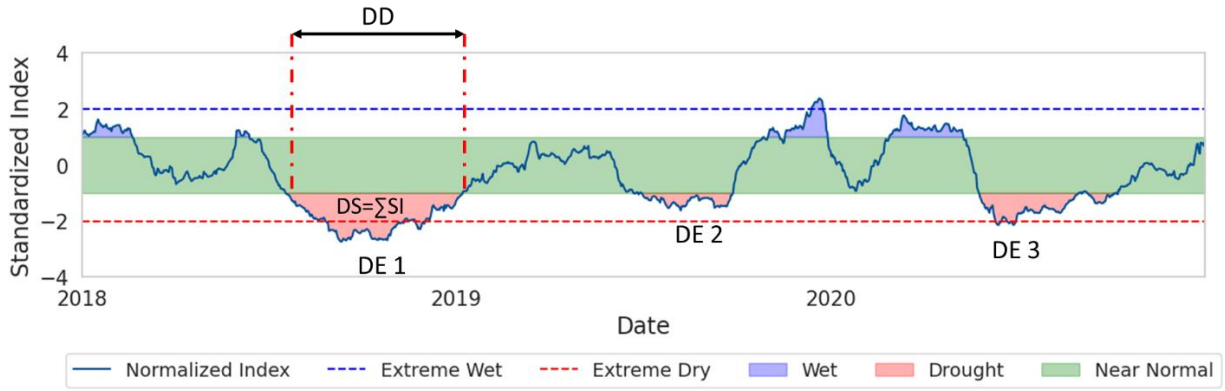


Figure 11: Drought characterizing using run theory. Where DE, DD, and DS refer to Drought episodes, Duration, and Severity respectively.

DS and DD were computed using equations 8 and 9 respectively.

$$\mathrm{DS} = \sum_{i=1}^{n} |SI_i| i = 1, \ldots, n \qquad \text{equation 8}$$

$$\mathrm{DI} = \frac{DS}{DD} \qquad \text{equation 9}$$

Where SI_i is the standardized drought index of interest, in this case, SPEI, SSWI, or SGWI, scored from i-time index to n.

3.2.3 Vegetation Condition Indices Computation

In this sub-section, two vegetation condition indices were used to monitor the dynamics of the distinct vegetation properties during drought conditions. NDWI monitors the plant hydrology conditions by utilizing the high reflectance from the cell mesophyll and water absorptions at NIR_{860} and $SWIR_{1240}$ wavelengths respectively (B. C. Gao, 1996). Conversely, NIRv quantifies the attribution of vegetation on the pixel reflectance by computing a multiple of the Normalized Difference Vegetation Index (NDVI) and NIR band reflectance (Zeng et al., 2022). NDVI refers to the normalized difference between the Red_{640} and NIR_{860} bands. Figure 12 shows the reflectance curves of vegetation and soil across the wavelengths of interest.

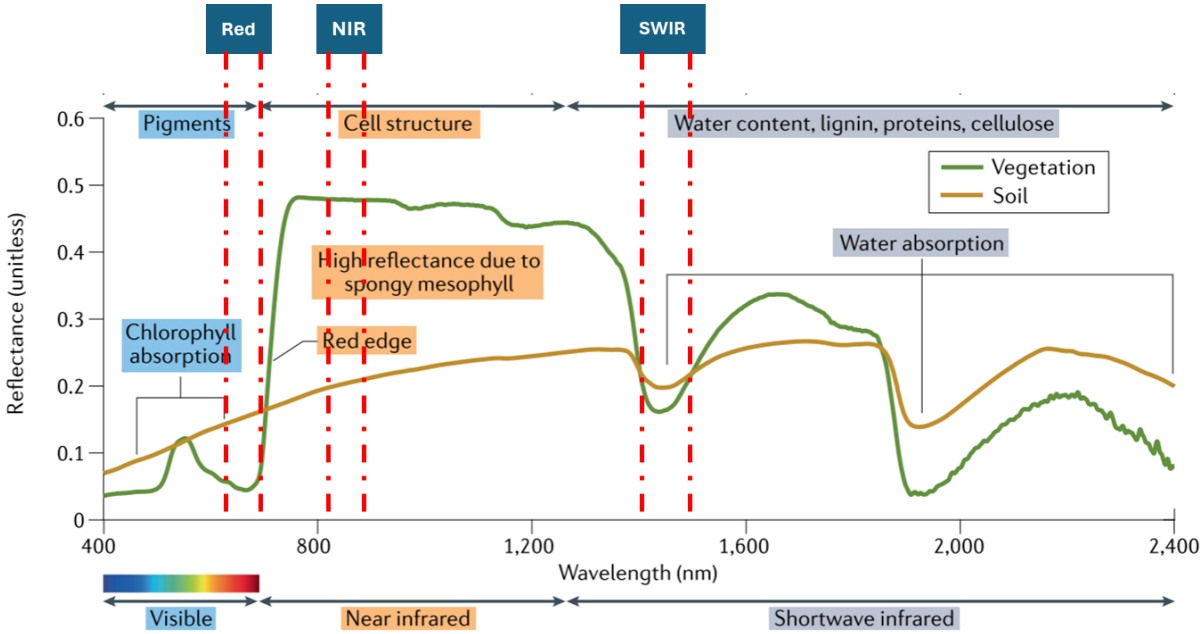


Figure 12: Reflectance behavior of vegetation and soil across wavelengths of interest. Diagram edited from the original version by Zeng et al. (2022).

From the extensive collection of optical vegetation indices, NIRv and NDWI were selected to be used in this study due to their proven robustness to various sources of uncertainty in optical vegetation monitoring (Zeng et al., 2022). NIRv enhances the impacts of vegetation on the NDVI by multiplying with NIRv to mask the contributions of the background noises whereas the use of long wavelength SWIR makes NDWI less sensitive to clouds compared to other plant biochemical indices such as red-edge Normalized Difference Vegetation Index (NDVIre). Moreover, Leaf Area Index (LAI), NDVI, and Enhanced Vegetation Index (EVI) were used for the development of 1Km datasets thus they won't be used in this study.

Figure 13, demonstrates the steps taken to extract vegetation condition anomalies from the broadband remote sensing dataset, MOD09GA latest version 6.1. Initially, MOD09GA was retrieved and processed in the Google Earth Engine code editor, cloud platform. The image processing procedures i.e., clouds and cloud shadow masking as well as no-data pixel values treatment explained in detail in the data processing section was the first procedure before indices computation. Additionally, spatial upscaling and temporal aggregation of the daily, are also explained in detail in the data processing stage. Below the methods used to compute vegetation indices from the MOD09GA bands are presented.

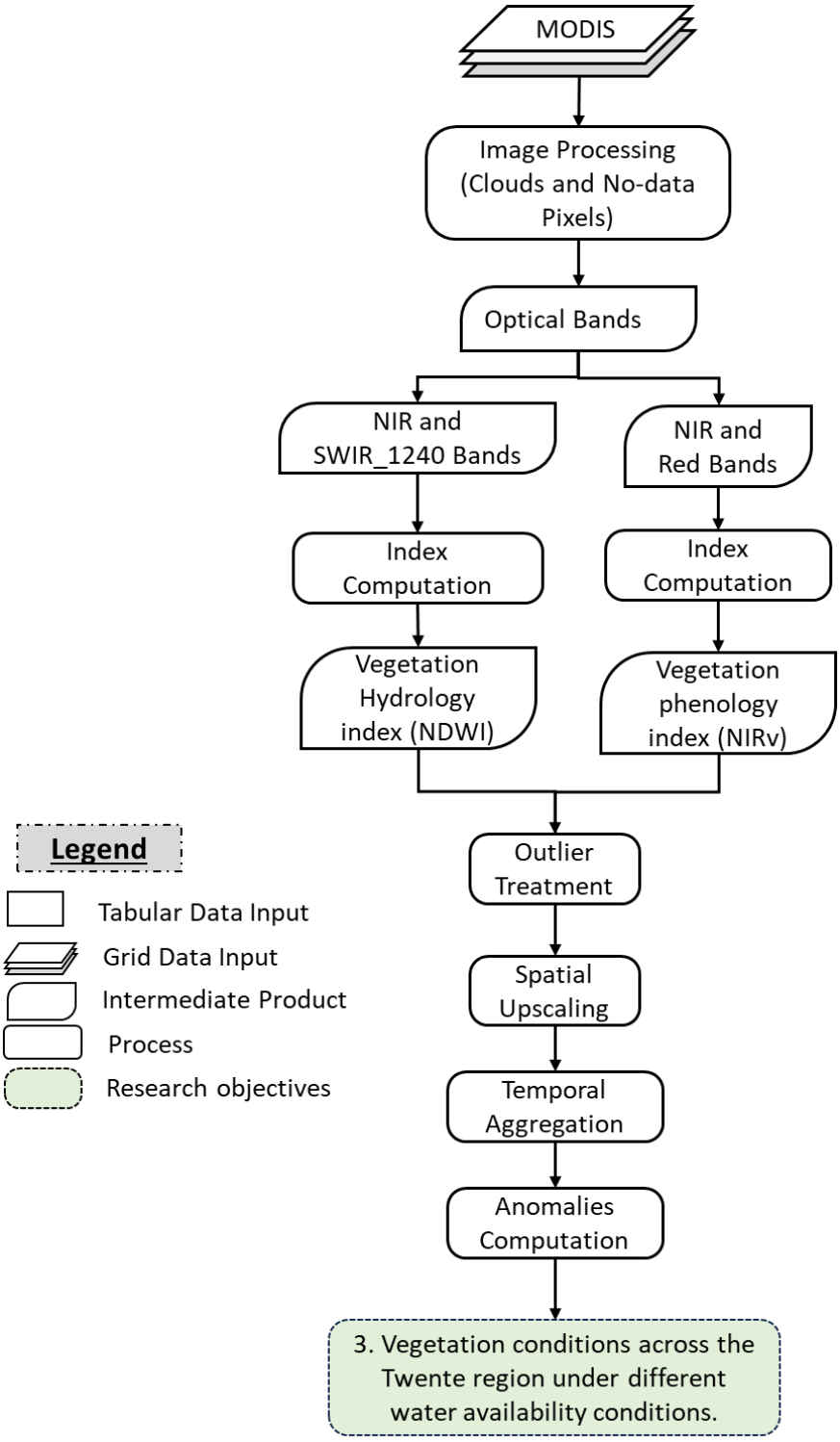


Figure 13: Demonstration of the methodological approach adopted for the third objective.

3.2.3.1 Vegetation Condition and Anomalies Computation

NIRv and NDWI were computed using MOD09GA version 6.1 bands as shown in equations 10 and 11 respectively:

$$NIR_v = \frac{NIR-Red}{NIR+Red} * NIR \quad \text{equation 10}$$

$$NDWI = \frac{NIR-SWIR_{1240}}{NIR+SWIR_{1240}} \quad \text{equation 11}$$

Where $SWIR_{1240}$ NIR and Red are MOD09GA reflectance bands 5, 2, and 1 respectively, NIRv is the Near Infrared reflectance of terrestrial Vegetation index. Finally, a monthly time series of vegetation indices was generated by computing an average value of both NIRv and NDWI throughout the respective month.

Equation 12 was used to derive the vegetation indices anomaly (Veg_{anom}) for each location. A simple vegetation anomaly representing the deviation of a vegetation index (Veg_i) from the mean pixel value over the 20-year study period.

$$Veg_{anom} = \frac{Veg_i - Veg_{\propto}}{Veg_{\sigma}} \quad \text{equation 12}$$

Where Veg_i is an incident vegetation index of interest (NIRv or NDWI) at a certain point in time and Veg_{σ} and $Veg_{\propto}$ represent the mean and standard deviation of the respective vegetation index at a pixel level.

After the extraction of the vegetation indices time series, preprocessing was necessary to ensure a clean vegetation indices time series for the computation of vegetation anomalies. In addition to the clouds and No data pixel cleaning explained in section 3.1.1, time series visualizations (line plots) were used to detect extreme and abnormal vegetation indices values. The sensitivity of NDWI to noises such as from soil water and snow cover absorption impacts was also addressed during NDWI computations (B. C. Gao, 1996; Zeng et al., 2022). Enhanced absorbance by soil water and snow cover at $SWIR_{1240}$ bandwidth leads to an extremely positive difference between $SWIR_{1240}$ and NIR, especially in scattered vegetated areas. For anomaly computations, outliers in the time series can significantly alter the interpretations of drought intensity (Rutger Weijers, 2020). To prevent corrupting the vegetation index time series, in this study, extremely negative and positive NDWI and NIRv values were substituted by the 97.5th and 2.5th percentiles of the pixels' time series (Example Appendix Figure S4).

3.2.4 Causal Effect Analysis with Convergence Cross Mapping (CCM)

In the current objective, a causal-effect analysis was performed to assess the causation interactions between drought and vegetation anomalies computed in the Witharen groundwater region for the period spanning 2000-2020. Figure 14 provides an overview of the approach adopted in this objective. Meteorological, soil moisture, and groundwater drought anomalies were acquired from objective 2 outcomes while from the third objective NDWI and NIRv anomalies were retrieved. Both indices were temporally matched into a monthly scale time series yielding on average 220 matchups in each location.

Recent studies explore a range of Causal Inference Methods (CIMs) including but not limited to graph-based algorithms, Ganger causality, Transfer Entropy, and Convergent Cross Mapping (CCM) to infer causalities in complex systems, particularly in the field of hydrometeorology (Bonotto et al., 2022; Ombadi et al., 2020; Z. Zhang et al., 2023). For instance, Delforge et al. (2022) emphasized the suitability of CIMs to decipher non-linear causations. In light of these developments, the selection of non-linear CCM to infer drought causalities in the current study is well-founded. Moreover, the demonstrated capacity to detect

physically explainable hydrometeorological causal interactions on a local scale motivated the selection of non-linear CCM to infer drought causalities in the current study (Ombadi et al., 2020).

CCM was designed to quantify the predictability of one variable (Y) in a time series to infer the causal influence of another variable (X) in a series (Sugihara et al., 2012). CCM assumes that if X causes Y, then in a historical time series of affected variable Y, the causal variable X leaves information signatures to influence the affected variable. For that case, by using Takens' theorem (Time-delayed embedding) the causal variable was estimated from the signature information in the affected variable. The process is also called cross-mapping and the information is referred to as shadow manifolds (Ionita et al., 2022).

Cross-mapping skill (ϱ) was computed as the correlation between the observed and cross-mapped X time series to quantify the cross-mapping accuracy. The actual causal interaction of X on Y was then distinguished from spurious causation if the ϱ score increases with the increase in library size, hence convergence (Bonotto et al., 2022). CCM causal inferring model can be summarized using equation 13:

$$X_{t+1} = X_t(r_x - r_x X_t - \beta_{x,y} Y_t) \quad \text{equation 13}$$

Where r_x is the variable X intrinsic growth rate, $\beta_{x,y}$ is the quantified influence of X on Y dynamics, and t and t+1 are the consecutive time steps of the time series dynamics. To estimate the causal influence of Y on X the equation can be reversed.

A detailed explanation of how CCM establishes causation can be referred to in a study by Sugihara et al. (2012).

In this study, CCM causal analysis was implemented in the R environment using the 'multispatialCCM' package. A uniform time lag ($\tau = 1$), and the best embedding dimension (E) were obtained as the one that produces the best ϱ output were the specified CCM parameters. For each CCM interaction modeling, a default 100 number bootstrap iterations were simulated to obtain median ϱ, a representative of the CCM causal results for the particular location.

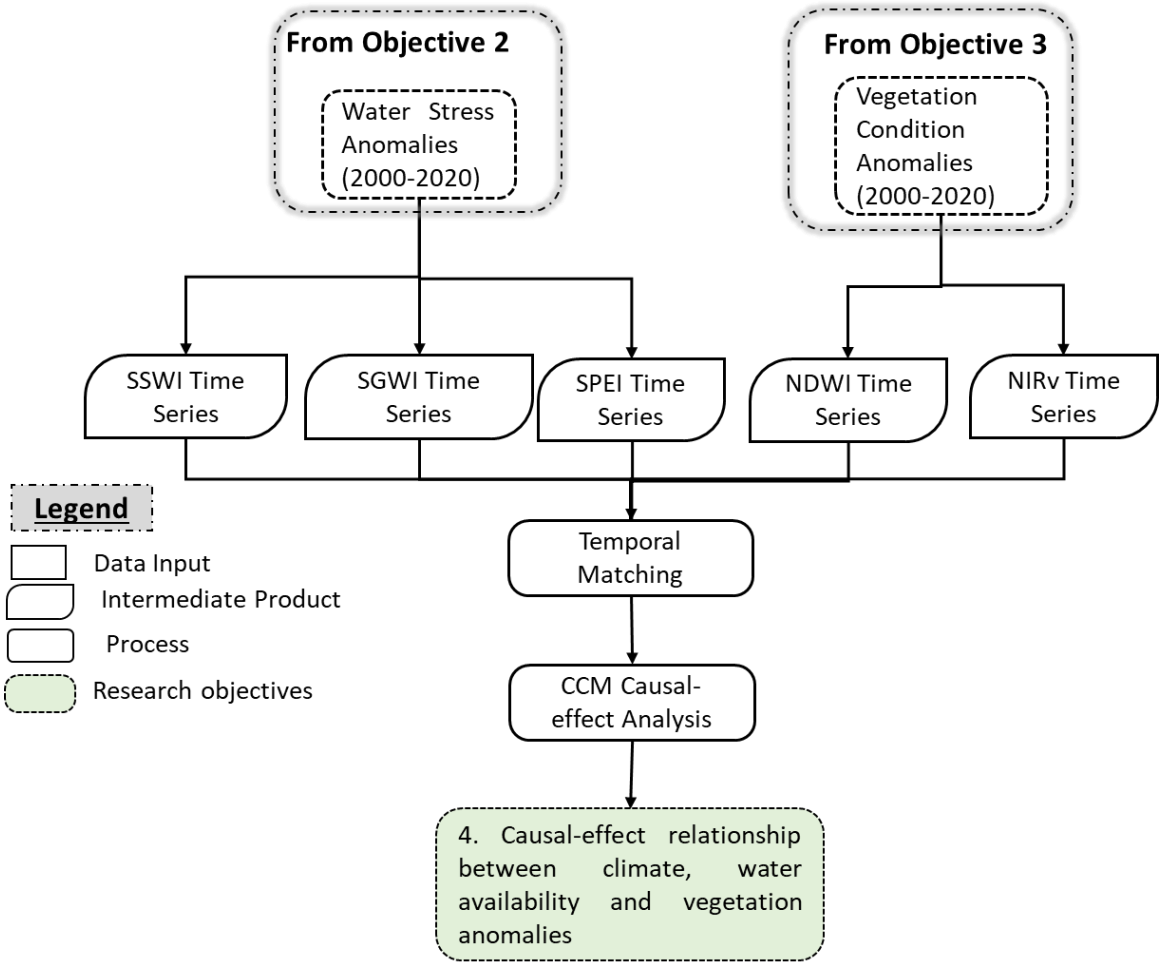


Figure 14: Demonstration of the methodological approach adopted for Objective 3

4 RESULTS

This chapter provides an overview of the results of the application of the methodologies explained in the previous chapter. It is structured into four sub-sections, each corresponding to a specific objective that collectively contributes to unraveling the causal dynamics between drought and vegetation anomalies.

- The first sub-section delves into the results and evaluation of the two Root Zone Soil Moisture (RZSM) estimation methods.
- The second sub-section explores the estimated time series of the three drought anomalies (Meteorological, Soil moisture, and groundwater drought)
- The third sub-section examines the time series couplings and the correlations between drought anomalies and vegetation anomalies.
- Finally, this chapter concludes with a detailed analysis of the Convergent Cross Mapping (CCM) causal inferences detected between vegetation conditions and drought anomalies.

4.1 RZSM Estimation Methods.

In this subsection, an overview of the results of RZSM estimations by the two evaluated methods is presented. In the first part, the optimal model parameters for the evaluated methods ELPF and CDF matching method are examined. Thereafter, the performance of the two models was evaluated using 4 accuracy metrics, Root Mean Square Error (RMSE), Bias term, correlation coefficient (r), and the unbiased Root Mean Square Error (ubRMSE). In the concluding subsection, the time series examination of the most suitable method for RZSM estimation based on the evaluated accuracy metrics is presented.

4.1.1 Optimal RZSM Model Parameters.

a) Exponential Low Pass Filter (ELPF).

In Table 4, the optimal delayed time required for the 1KmSSM signals to match in-situ soil moisture signals at 10cm, 20cm, and 40cm are presented for the selected soil moisture monitoring stations. Generally, there is an increase in the optimal T value with the depth of the target in-situ soil moisture signals. For the 7 evaluated stations, the median optimal delayed T value was 16, 17, and 33 days for 10 cm, 20 cm, and 40 cm depths respectively. Indicating an immediate response to surface soil moisture dynamics at the shallow soil depths (10 and 20 cm) compared to the deeper layers (40 cm).

Interestingly, some stations had an opposite trend in the optimal delayed time respective to the matching depth. For instance, stations ITC_SM07, 09, and 15 required longer delayed signals to match 10 cm soil moisture compared to the 20 cm soil moisture signal. ITC_SM09 located by the river Buursebeek had the highest difference of 12 days between the optimal T value required to match 10cm and 20cm soil moisture dynamics. The soil profile of the ITC_SM09 station is composed of a loamy texture at the topsoil with fine sand in the sub-surface layer. A significantly high resistance to vertical water movement due to the compaction of top loamy textured soil might explain a highly dynamic 20 cm soil moisture dynamic compared to the top layer (10 cm).

Moreover, the analysis reveals significant variability in the mean optimal delayed time for the three depths among the seven evaluated stations. Specifically, stations ITC_SM03, 07, and 15 had an average optimal T parameter above 25 days while ITC_SM09 and 18 required less than 11 days, the lowest among the evaluated

stations. The difference in optimal T parameter among stations further the difference in coupling nature between 1KmSSM and in-situ deeper soil moisture content time series across the study area.

Table 4: Optimal delayed time in days for the selected stations and 10, 20, and 40cm depths

Station ID	10 CM (DAYS)	20 CM (DAYS)	40 CM (DAYS)
ITC_SM02	5	17	27
ITC_SM03	34	44	69
ITC_SM07	19	12	48
ITC_SM09	14	2	15
ITC_SM13	16	18	30
ITC_SM15	23	20	35
ITC_SM18	2	6	12

b) CDF Matching

The optimal polynomial fit that produces the most accurate RZSM at 40 cm from 1KmSSM is presented in Table 5. From these outputs, it is evident that in most of the soil moisture stations, a complex polynomial line was required to model the relationship between 1KmSSM dynamics and the respective difference with RZSM incidences. Specifically, at least an optimal degree of 3 was required for each station to achieve maximum accuracy in predicting Δ from 1KmSSM dynamics. 4 monitoring stations out of the 7 evaluated achieved a maximum accuracy using a 5th order polynomial model while only the ITC_SM03 station required the 6th order polynomial model fitting suggesting a more complex relationship.

Table 5: Optimal polynomial fit for each station.

Station (ITC_)	SM01	SM03	SM07	SM09	SM13	SM15	SM18
Polynomial Fit	5	6	5	5	3	3	5

4.1.2 RZSM Accuracy Assessment

Figure 15 presents the distribution summary of 4 accuracy metrics used to evaluate the predicted $RZSM_{ELPF}$ and $RZSM_{CDF}$ estimated using ELPF and CDF matching methods respectively. In general, both methods were able to reproduce the time series dynamics of the in-situ RZSM, consistently scoring a satisfactory correlation coefficient of above 0.73. A strong linear correlation scored by both $RZSM_{ELPF}$ and $RZSM_{CDF}$ indicates a reliable ability of the two models to infer the time series dynamic of RZSM from the 1KmSSM time series. However, the ubRMSE range between 0.04 m3/m3 to 0.08 m3/m3 suggests an alarming performance in some locations.

Between methods, $RZSM_{ELPF}$ consistently scored more desirable accuracy metric scores as compared to $RZSM_{CDF}$. A median Pearson correlation coefficient of 0.87 scored by $RZSM_{ELPF}$ outperformed $RZSM_{CDF}$ by 0.08 (Plot 15 B). Moreover, similar to the correlation coefficient results, Plot 15 A reveal that, $RZSM_{ELPF}$ outperformed $RZSM_{CDF}$ by an average RMSE difference of 0.002 m^3/m^3. Station ITC_SM07 scored an RMSE of 0.08 m^3/m^3 between 1KmSSM and in-situ soil moisture. This station consistently produced the

highest residual measure after the application of both ELPF and CDF matching methods. This indicates the impact of the accuracy of the 1KmSSM dataset on the RZSM estimation results.

Bias estimation results indicate that the two methods constantly overestimated RZSM (Plot 15 D). Specifically, in at least 5 of the 7 evaluated stations, both $RZSM_{ELPF}$ and $RZSM_{CDF}$ exhibited positive biases. An outstanding result was observed in station ITC_SM07 in which $RZSM_{ELPF}$ dramatically underestimated RZSM, reaching a bias of -0.095 m^3/m^3 leading to an RMSE of 0.12 m^3/m^3, the highest to be recorded by the approaches.

Moreover, improved ubRMSE evaluation compared to RMSE results indicated that most of the residuals on the estimated RZSM by both ELPF and CDF methods are mostly contributed by bias and less by error. Specifically, plot 15 C reveals that $RZSM_{ELPF}$ had an average improvement of 0.016 m^3/m^3 while for $RZSM_{CDF}$ the RMSE improved by an average of 0.007 m^3/m^3. Moreover, $RZSM_{ELPF}$ for the station ITC_SM07 produced an ubRMSE of 0.072 m^3/m^3, a significant improvement from the RMSE of 0.12 m^3/m^3, further indicating a significant contribution of biases on the computed RMSE.

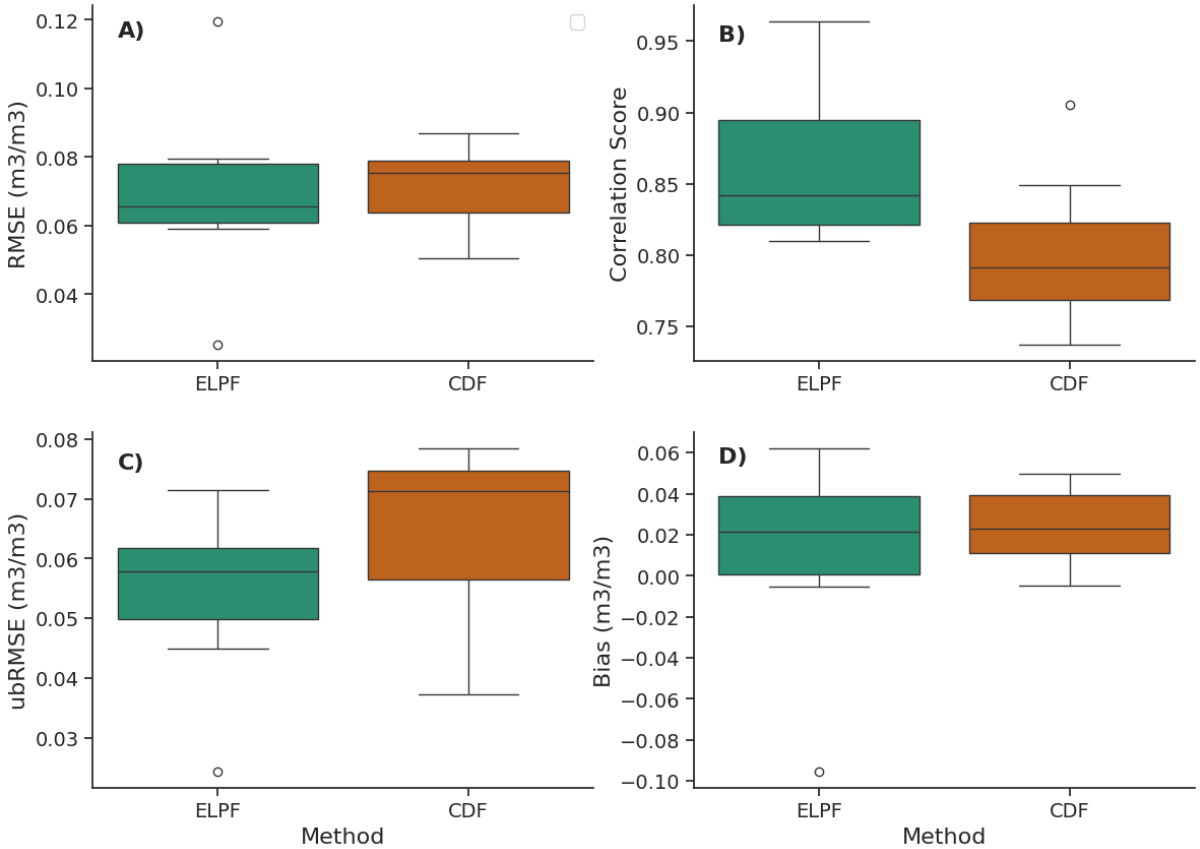


Figure 15: Distribution of accuracy metrics scored by the two evaluated methods. A. RMSE, B. Correlation Score, C. ubRMSE, and Bias term.

4.1.3 Best Performing Method

In this sub-section, the time series of the estimated RZSM using the best-selected method compared to the in-situ RZSM, is presented. From the accuracy assessment sub-section, it is evident that $RZSM_{ELPF}$ consistently outperformed $RZSM_{CDF}$ in both correlation and residual assessment. Additionally, the

complexity of the lines fitted by polynomial degrees 3, 5, and 6 complicates the intuitive interpretation of the relationship between SSM and Δ. For that case, a median optimal T value of 16, 17, and 33 days for 10 cm, 20 cm, and 40 cm depths respectively was used to depth filter 1KmSSM time series before scaling to RZSM estimates.

The application of median optimal T values on the calibrated stations yields a mean correlation coefficient and RMSE of 0.86 and 0.07 m^3/m^3 respectively. Moreover, two uncalibrated stations ITC_SM08 and 14 were tested and both scored an RMSE score less than 0.07 m^3/m^3 and a correlation coefficient higher than 0.81. This indicates the potential of the ELPF method transferability and generalizability. However, extremely low 1KmSSM accuracy and missing data in stations such as Hupsel and Twente Airport limited the testing across other uncalibrated Twente stations

Figure 16 demonstrates the $RZSM_{ELPF}$ time series covering the validation period of the three selected stations of which two (A, C) were not involved in the original experiment to calibrate the optimal T value. From the time series, the predicted $RZSM_{ELPF}$ was able to match most of the in-situ measurements of RZSM peaks during the validation period. However, steep rises in the in-situ soil moisture for instance in March 2020 of plots 8, A, and B were not matched by the $RZSM_{ELPF}$ time series. A similar pattern was observed in July 2019 and 2020 of the same plots. On the other hand, the smooth peak observed in March 2020 for plot 6, C was perfectly matched by the $RZSM_{ELPF}$ time series.

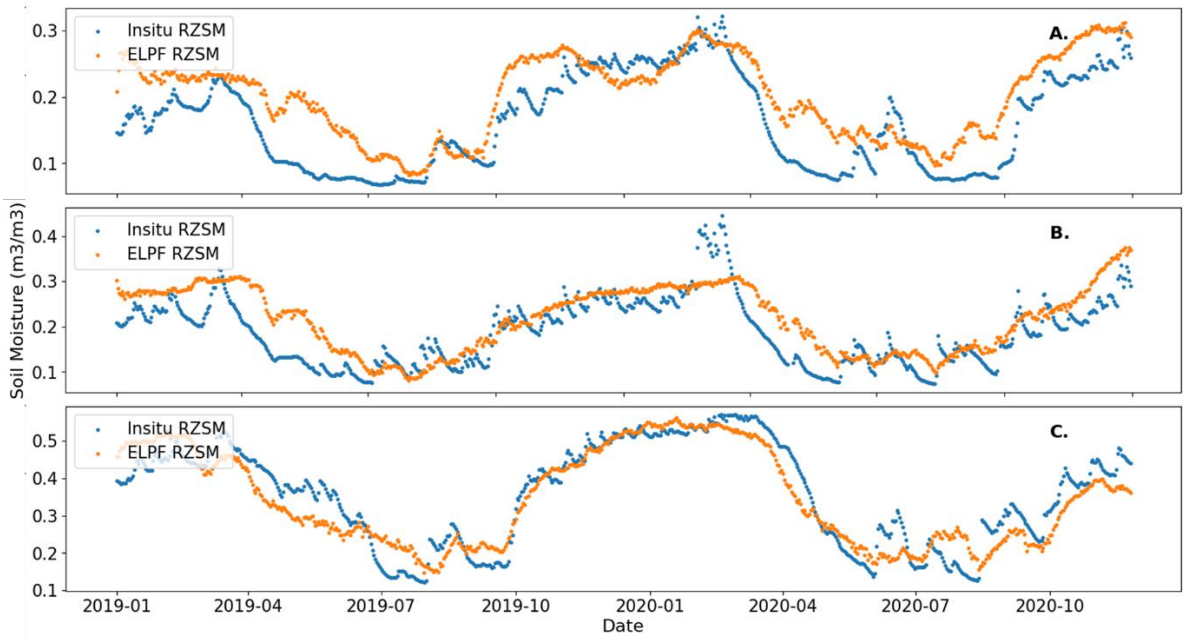


Figure 16: $RZSM_{ELPF}$ vs In-situ RZSM time series for stations ITC_SM08, 13, and 14 (A, B, and C respectively).

4.2 Drought Anomalies and Drought Characterization:

In this part, the temporal drought evolution of the 27 Witharen groundwater monitoring stations for the period spanning 2000 to 2020 is presented. Three drought indices, i.e., SPEI-1, SSWI-1, and SGWI-1 were computed for the monitoring of meteorological, soil moisture, and groundwater drought conditions. Validated ERA5's precipitation and actual evapotranspiration, $RZSM_{ELPF,}$ and in-situ point scale groundwater measurements were used as input for the computation of the three indices respectively. The calibrated ELPF parameters from the first objective were transferred and used for the estimation of the RZSM time series in the Witharen region.

Figure 17 provided the temporal evolution of the mean meteorological, soil moisture, and groundwater drought indices, as assessed using SPEI-1 together with the mean SGWI-1 and SSWI-1 score to provide a general overview of the drought anomalies in the study area. The time series of the drought indices indicated a variation in the onset and offset of the mean drought events among the three indices. Typically, drought events are initiated by the SPEI-1 anomalies before proceeding to the mean anomalies of both SSWI-1 and SGWI-1 indices. Episodes such as wet anomalies in 2007, 2010, and 2014, as well as dry episodes in 2011 and 2020, demonstrate this phenomenon. During those episodes, SPEI-1 experienced more pronounced positive or negative anomalies together with a quick recovery to a normal condition compared to SSWI-1 and SGWI-1. This observation was expected as it indicates the general propagation of the natural drought, initiated by precipitation scarcity before propagated to soil water anomalies.

Contrary to the general trend, various seasons in the time series of the drought indices revealed developments that contradict the expected general flow of natural drought. Observations of the dry episodes in 2003 and the wet episodes in 2004 indicated the fall and rise of SSWI-1 and SGWI-1 anomalies that not only proceeded but also outscored, and outlasted the severity measured by the SPEI-1 index. Additionally, the time series observations indicated the seasonal decoupling of water anomalies. For instance, dry episode detected by SPEI-1 and SGWI-1 during the winter of 2006 was matched with near-wet anomalies of SSWI-1. These observations suggest the complexities in drought association analysis and the impacts of other hydrometeorological factors on the occurrence of water storage anomalies apart from meteorological.

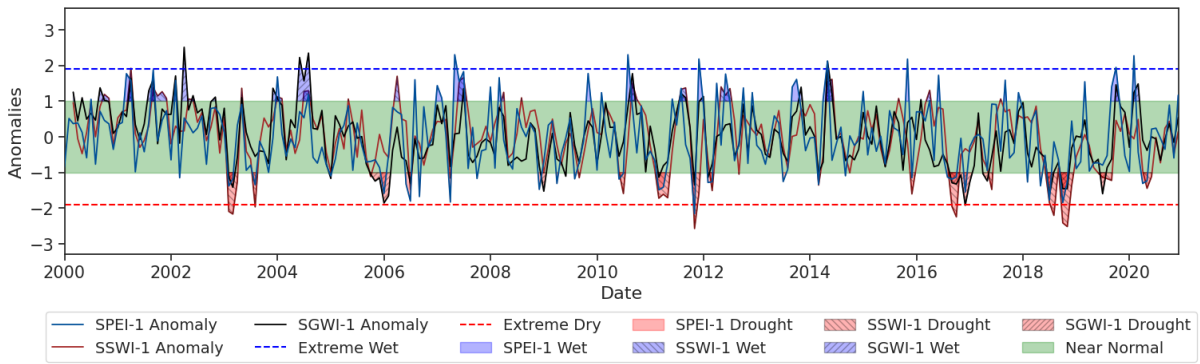


Figure 17: Temporal pattern of Meteorological, Soil Moisture, and Groundwater drought indices as monitored using SPEI-1, SSWI-1, and SGWI-1 indices respectively for a study period spanning 2000-2020.

Figure 18 presents the frequency, duration, and intensity of drought events identified using the mean study area drought indices. Within the study timeframe, SPEI-1 identified the highest frequency of dry episodes (35) while 21 and 17 episodes were identified by SSWI-1 and SGWI-1 respectively. Most of the drought episodes were identified during the last 10 years of the study period, such that, 65% of the SSWI-1 and SGWI-1 dry anomalies were detected while 54% of SPEI-1 drought events were detected in the same period. The observed intermittent nature of meteorological drought, marked by the periodic recovery of some episodes contributes to the high number of drought events. For instance, 4 evanescent meteorological droughts were identified from February to November 2018 with brief recoveries in April, May, August, and September.

The intensity of drought episodes varied depending on the type of index of interest. Specifically, SPEI-1 identified the 2011 drought episode as the most intense meteorological drought with the highest DI score of 2.15 while SSWI-1 and SGWI-1 scored the highest DI values of 2.1 and 1.6 during the 2018 and 2019 drought episodes respectively. A progressive increase in SPEI-1, SSWI-1, and SGWI-1 intensity was observed between 2017-2019, during which neither of the drought indices recovered to a wet anomaly during that period suggesting a propagation of drought effects between episodes (figure 17). Additionally, the average DI across the study period stood at 1.23, 1.44, and 1.27 for the drought indices SPEI-1, SSWI-1, and SGWI-1 respectively suggesting soil moisture experienced the most intense drought anomalies compared to groundwater and precipitation.

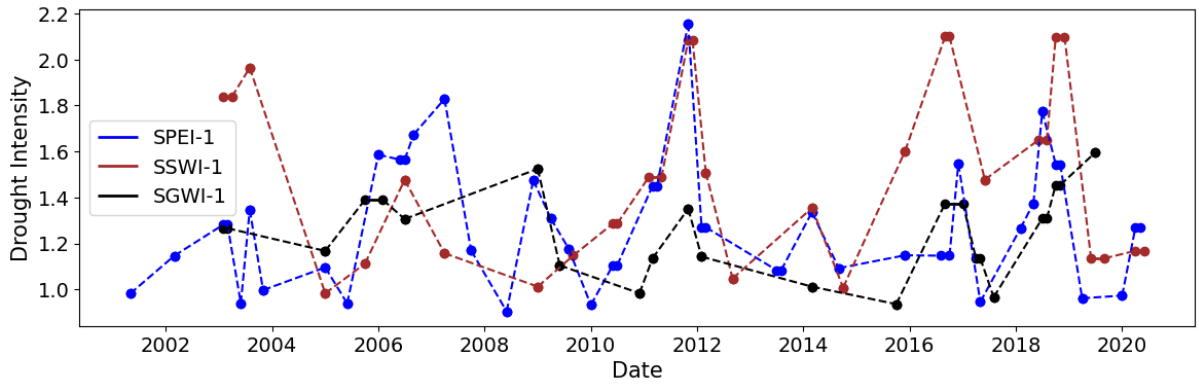


Figure 18: Change of Drought Intensity (DI) score across various identified drought episodes.

In examining the relationship at individual monitoring locations, temporal coupling was quantified using linear correlation coefficient scores. Figure 19 provides an overview distribution of the linear correlation scores between drought index pairs. The analysis reveals a significant positive linear relation between the three pairs of drought indices. Notably, SSWI-1 association with SGWI-1 scored a median correlation coefficient of 0.56, the highest between the pairs. This was closely followed by the other two pairs with median scores of 0.54 and 0.47. A more consistent correlation score among monitoring locations was observed in the SPEI-1 and SSWI-1 linear coupling while the SSWI-1 and SGWI-1 pairs showed much-elongated boxplots, suggesting a variable linear correlation.

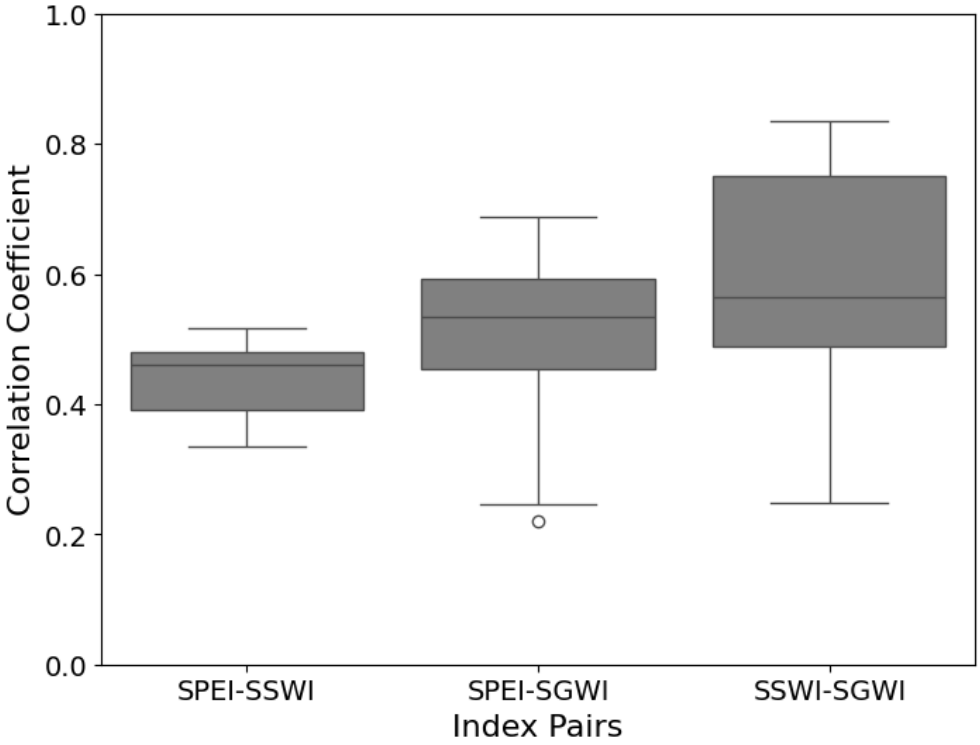


Figure 19: Distribution of linear correlation coefficient between drought indices.

To assess the impacts of spatial variables on the association between drought anomalies, the linear correlation coefficient scores were analyzed against various factors that might influence the linear coupling between drought indices. Dominant land cover properties, soil texture, and water table depth were all considered. A heterogenous land use activity at a 1 km soil moisture pixel meant it was difficult to assign an exclusive land cover property to an individual pixel. Moreover, only 4 distinct sandy soil classes were identified among the monitoring locations based on the Bofek 2020 with an overlapping correlation score between indices meant that it was not enough to explain the variation.

A notable association was found between the mean groundwater depth of the location with the occurrence of groundwater drought anomalies. As highlighted in the data exploration section, the study area is characterized by a shallow aquifer with only two wells deeper than 2.7 m and 3 under 0.75 m with the remaining wells falling between 0.89 m to 1.3 m. The comparisons between the deepest and shallowest monitoring locations indicated that shallow groundwater aquifers are more linearly coupled with surface water dynamics. Specifically, SGWI-1 of the deepest aquifer locations (B22C0129 and B22C0131) scored the lowest correlation coefficient with SPEI-1 (<0.25) while B22C0134 and B22C0107 both scored above 0.6.

However, this pattern weakened when considering correlation scores between SSWI-1 and SGWI-1. The highest correlation between SSWI-1 and SGWI-1 (>0.70) was scored by B22C0551, B22C0558, B22C0559, and B22C0068 wells whose depth range between 0.7 m to 1.4 m. Two wells B22C0086 and B22C0132 which are 0.6 m and 1.3 m deep respectively both scored a correlation score below 0.3, the lowest SSWI-SGWI

correlation score. Whereas the identified shallowest wells B22C0134 and B22C0107 as well as the deepest well, B22C0129 scored a correlation coefficient ranging between 0.5 to 0.57.

Furthermore, Figures 20 A and B present the time series of the three drought indices for the monitoring well B22C0134 (the shallowest at 0.59 m) and B22C0129 (the deepest at 3.7 m) respectively. Revealed, a more stable B22C0129 groundwater anomaly time series compared to B22C0134. A total of 32 episodes were observed in the B22C0134 time series, 14 more episodes of negative SGWI anomaly when compared to B22C0129. Monthly negative SGWI-1 anomaly was endured in B22C0134 for 42 months whereas B22C0129 experienced only 38 months throughout the study period suggesting a more monthly propagated SGWI in B22C0129. The highest DI score by B22C0134 was 2.6 while 1.1 was scored by B22C0129.

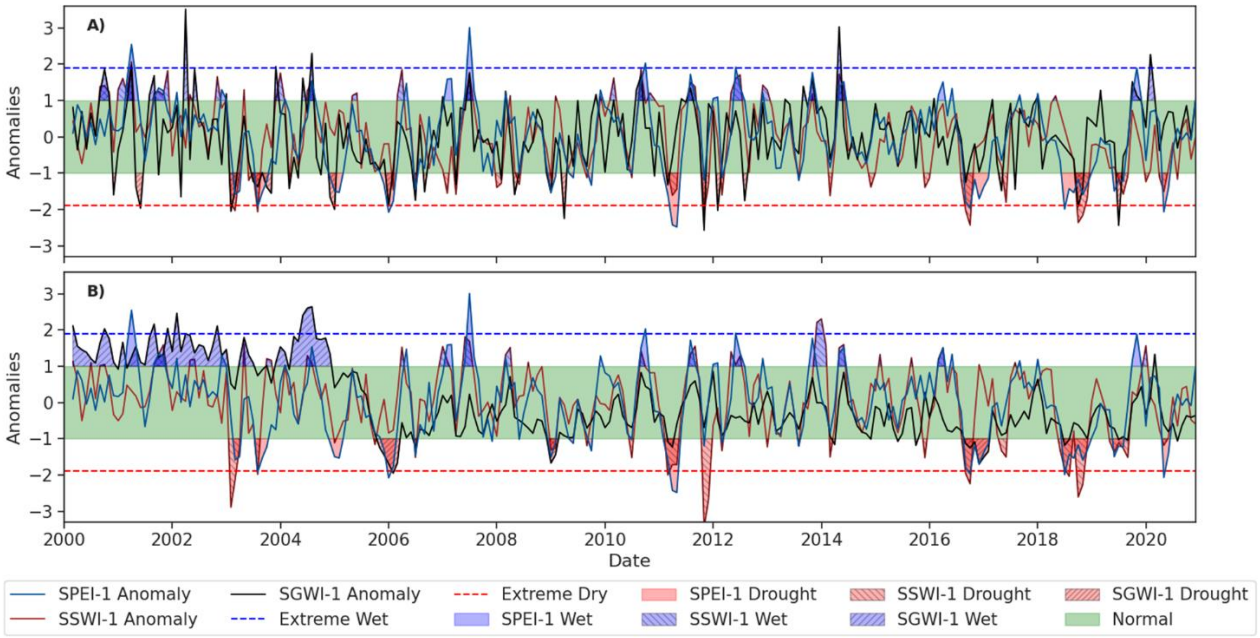


Figure 20: Temporal pattern of SPEI-1, SSWI-1, SGWI-1 for monitoring well A) B22C0134 and B) B22C0129

4.3 Vegetation and Drought Anomalies Association

In this objective, the association between drought anomalies (SPEI-1, SSWI-1, and SGWI-1) and vegetation anomalies (NDWI and NIRv) is illustrated. Initially, linear correlation and time series analysis between the vegetation anomalies were conducted to assess the coupling of the two indices. Furthermore, a zoom-out of the mean vegetation condition development during various drought episodes is presented to assess the decline and recovery of vegetation conditions during the on-set and off-set of drought episodes.

A correlation analysis between NDWI and NIRv was performed to evaluate the coupling between the vegetation indices in both the spatial and temporal dimensions. An overall linear correlation of 0.49 between the median NDWI and NIRv time series reveals a moderate linear relationship. Spatially, a correlation coefficient of 0.88 between the median NDWI and NIRv of the 27 used locations reveals a near-perfect agreement in the distribution of vegetation conditions. A notable deviation from the linear relationship was observed by station B22C0136 which had a moderate NDWI score despite scoring the lowest median NIRv (Supplementary figure S5).

Figure 21 presents the time series of the median NIRv and NDWI. A general agreement in the temporal trend of the vegetation anomalies can be observed with the synchronized rising and falling of the vegetation anomalies. Despite a near-perfect time series sync, the divergence can be observed at the onset of the negative anomalies. For instance, the dry anomalies of 2000, 2003, 2005, 2010, 2011, 2013, and 2014 were all preceded by NIRv-negative anomalies matched by an average of two months lag in NDWI anomaly before catching up with NIRv. During this, the trajectory of NIRv has started recovery to near normal condition. Another disparity was observed with the depth of the negative anomalies. During the drought episodes of the summer of 2003, 2018, 2019, and 2020 NDWI had a notably higher negative anomaly compared to NIRv.

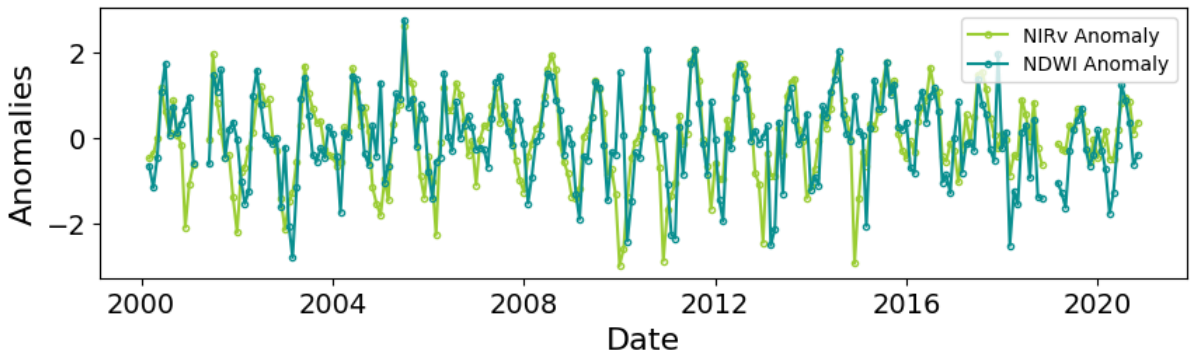


Figure 21: Time series of the median NDWI and NIRv in the study area

The linear correlation coefficient was also used to quantify the association between vegetation anomalies and drought indices (SPE-1, SSWI-1, and SGWI-1) (figure 22). Overall, linear correlation analysis revealed a generally positive but weak linear between vegetation anomalies and drought indices, indicating deterioration of vegetation conditions with drought intensity. Over the entire study period, a weak linear interaction was observed between anomalies of drought indices and NDWI while NIRv showed a majority of non-linear relationship with drought anomalies.

Furthermore, seasonal variations of the relationship were examined, adhering to the four seasons of the year as classified by KNMI (Siegmund, 2022). For the entire study area, the summer season showed the strongest positive relation between drought indices and vegetation anomalies with a mean coefficient of 0.21 by NIRv and 0.4 by NDWI followed by spring. In contrast, the winter season showed the weakest linear relation with both NDWI and NIRv scoring a mean non-linear correlation coefficient. This suggests a strong coupling between vegetation and water availability dynamics during the growing seasons (May to October) compared to the non-growing seasons (November to April)

Moreover, the difference in the correlation between the drought anomalies was also detected. NDWI had the highest mean correlation coefficient of 0.26 with SSWI-1 followed by SPEI-1 and then SGWI-1. A similar rank of linear relation was observed with NIRv correlation coefficient scores in the summer season. For NDWI, this rank was however different in Autumn, where a stronger linear correlation was scored with SSWI-1 (0.49) but followed by SGWI-1 and lastly SPEI-1. Whereas, with NIRv, SGWI-1 scored the highest linear correlation coefficient (0.2) during the winter season. Moreover, during the spring season, SPEI-1 consistently scored the highest linear correlation coefficient with the two vegetation anomaly indices.

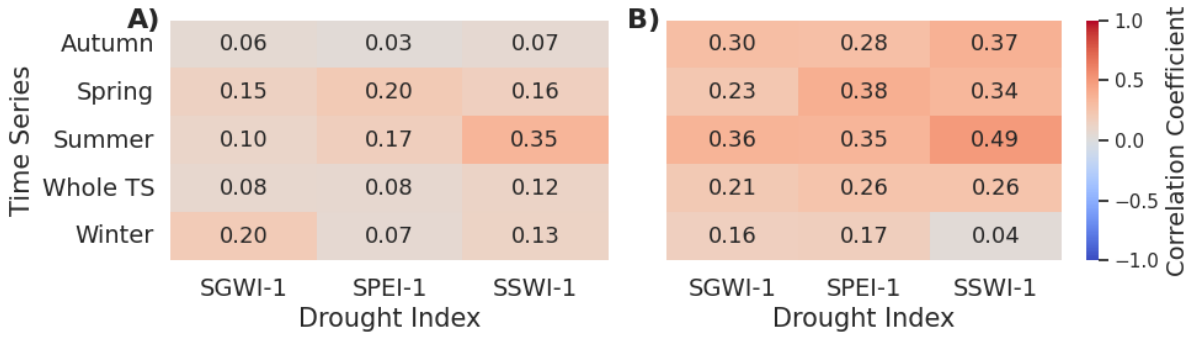


Figure 22: Pearson Correlation coefficient between drought and A) NIRv and B) NDWI anomaly. Whole TS refers to the whole time series of the study period.

Figure 23 presents the temporal development of the median vegetation conditions during various drought episodes. Coupling between vegetation anomalies and drought indices in summer, for example, the summer of 2003, 2006, and 2010 reflects on the observed positive temporal relationship (figure 22). However, focusing on the general time series, a complicated vegetation condition response to the general trend in the rising and falling of the monthly drought anomalies was observed.

Temporal lagging after the wet and dry anomalies of the drought indices is mostly observed in the time series. Drought episodes of summer 2003, 2006, and 2016, as well as wet episodes in the summer of 2010 and 2011, demonstrate this phenomenon. Periodic recoveries of vegetation conditions were also observed during the prolonged drought episodes of 2018 as well as that of the spring of 2011. The timing of vegetation anomaly changes relative to drought indices varied significantly, with delayed decline and recoveries of vegetation conditions observed in most cases.

Decoupling in the vegetation and drought anomalies was another phenomenon observed in the time series. For instance, less intense drought episodes of the summer of 2009 and 2013 were matched by the positive anomalies of both vegetation condition indices. Another example was observed during the winter of 2005, 2010, and 2011 in which wet drought anomalies were matched by negative vegetation anomalies by both

NIRv and NDWI. These temporal disparities in the vegetation response explain the observed weak linear correlation with the drought indices (figure 22).

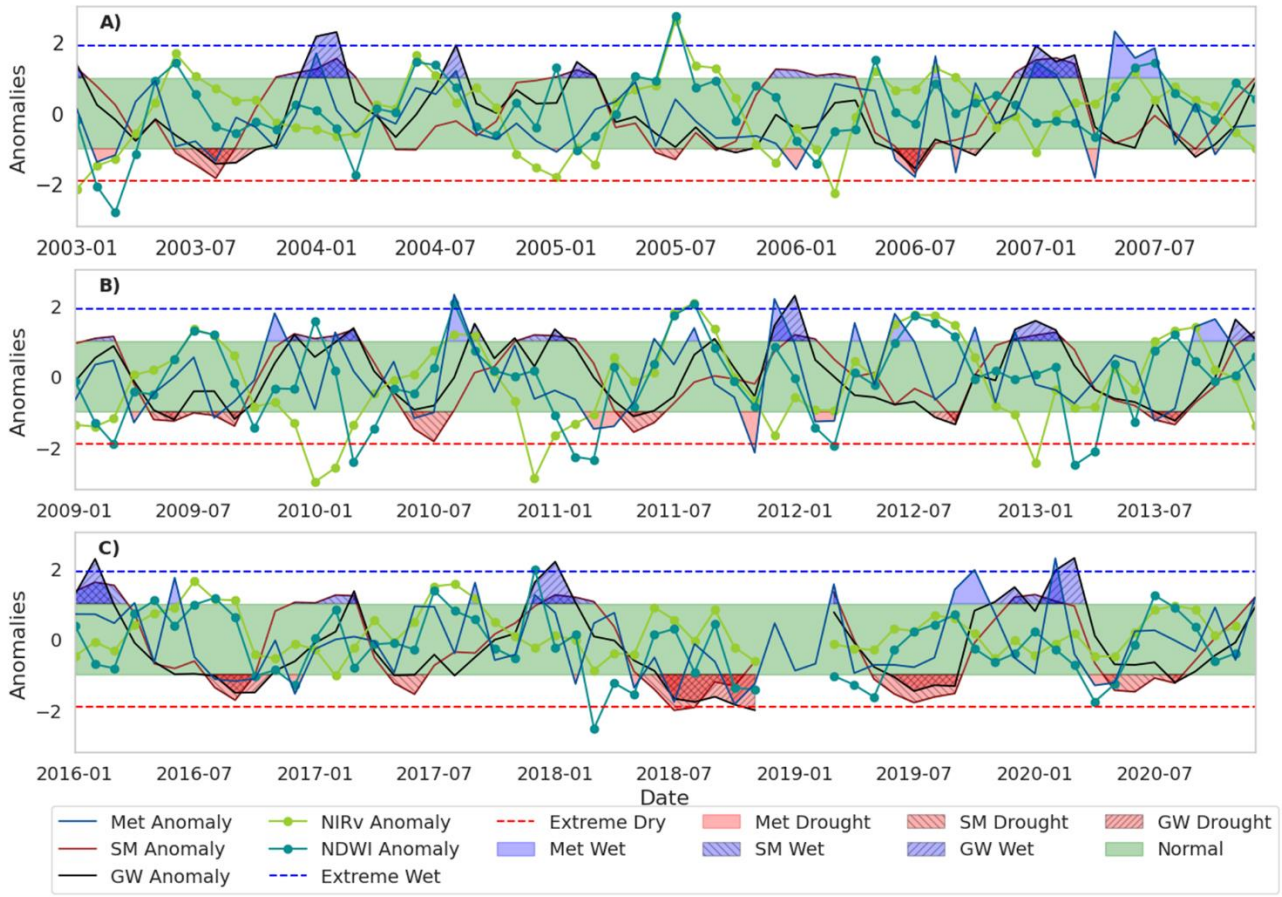


Figure 23: Vegetation condition development during intense drought episodes of a) 2003, 2006, b) 2011, and c) 2016, 2018, 2019, and 2020.

4.4 Drought Causal Analysis Results

In this section, the results of applying Convergent Cross-Mapping (CCM) for the analysis of causations in drought indices are presented. A total of 30 plausible drought causation interactions were evaluated using drought and vegetation indices computed in objectives 2 and 3 (flowchart in Figure 24). Additionally, the ability of Surface Soil Moisture anomaly (SSWI-1u) to capture drought causation interactions was tested in section 4.4.2. SSWI-1u was computed using the same approach as SSWI-1 (in section 3.2.2) but using SSM instead of the originally used RZSM. The reason for using SSWI-1u is to test if RZSM and SSM dynamics can provide different signals in drought causation analysis compared to RZSM.

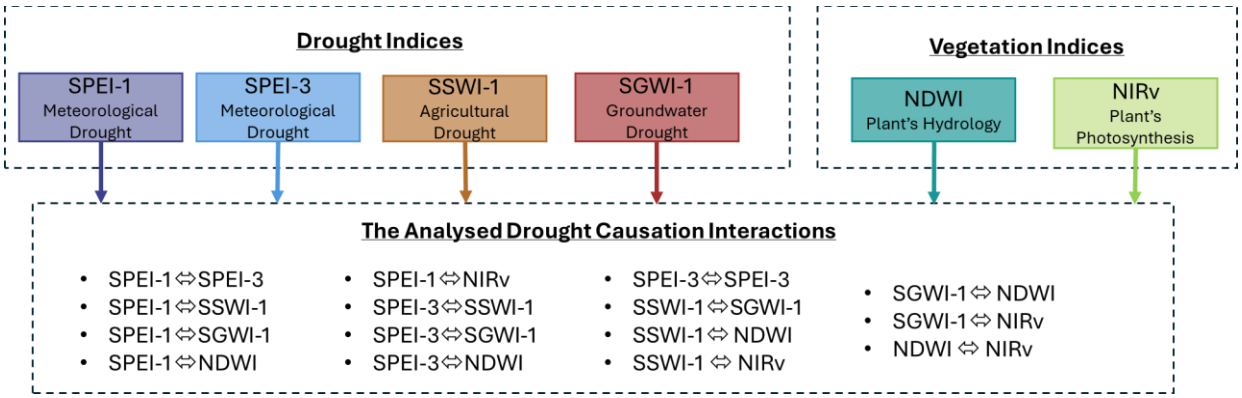


Figure 24: Flowchart of the drought causality analysis

Two main conditions were used to identify significant causation from spurious causality. First, is the convergence of the CCM, that is the cross-map skill (ϱ) increases with the library size (Sugihara et al., 2012). Second, non-spurious causation was determined by having a ϱ score at the maximum library size both significant at a 95% confidence level and higher than the linear correlation score between the time series pair (Ionita et al., 2022). In cases where significant cross-mapping skill was observed in both causation directions of the causation, the direction with the highest ϱ value at the maximum library size was identified as a dominant causal direction.

The interpretation of the causation results was based on what the index signifies at a particular scale. At a 1-month scale, drought indices identify the on-set and off-set of drought anomalies hence the causation results were associated with the initiation and decline of drought episodes. Both SPEI-1 and SPEI-3 were used in the current analysis to account for both the occurrence and propagation of meteorological drought (Rutger Weijers, 2020).

4.4.1 CCM Analysis Between Drought Anomalies

The results from Figure 25 provide a comprehensive overview of the CCM results in terms of the cross-mapping skill (ϱ) and linear correlation coefficient between pairs of drought and vegetation indices. The median ϱ among the evaluated 27 locations is presented by the solid line while the shaded area provides the upper and lower standard deviation boundaries. Hereafter, the median ϱ is used as the representative ϱ score in each causation direction.

Notably, the highest cross-mapping skill was observed between meteorological drought indices (from SPEI-1 to SPEI-3, ϱ=0.92) indicating a robust interconnection between meteorological drought initiation and accumulations. On the other hand, the lowest cross-mapping skill was observed in NDWI => SGWI-1 causal link (ϱ = 0.02), suggesting a minimal association between vegetation water content and groundwater anomalies.

Based on the explained conditions of significant causality, a total of 12 significant causal interactions were identified, of which 3 causal interactions were bidirectional. These bidirectional causalities were identified in Figure 25 A, L, and O between index pairs SPEI-1 and SPEI-3, SSWI-1 and NIRv, as well as NDWI and NIRv respectively. The SPEI-1=>SPEI-3 causal interaction had the highest ϱ than SPEI-3=>SPEI-1 indicating a stronger causal influence by SPEI-1 than the opposite causation direction. In the interaction with SSWI-1, NIRv demonstrated a weaker causality influence compared to the opposite direction whereas, with NDWI, NIRv scored the highest ϱ hence the strongest causal influence.

Additionally, the analysis identified significant unidirectional causalities particularly from SPEI-1 to SSWI-1, SPEI-3 to SSWI-1, SSWI-1 to SGWI-1, SSWI-1 to NDWI, SGWI-1 to NDWI and SGWI-1 to NIRv as shown on Figure 25 B, F, J, K, M, and N respectively. The opposite direction of these causal interactions either didn't converge or scored a lower cross-mapping skill at a maximum library size. For instance, causal directions from NDWI to SSWI-1, NDWI to SGWI-1, and NIRv to SGWI-1 did not converge significantly, hence they were considered insignificant causal interactions. For the causation directions from SSWI-1 to SPEI-3 and SPEI-1 there was convergence, however, the ϱ at the maximum library size was lower than their respective linear correlation coefficients suggesting a weak or spurious causation.

On the other hand, several anomaly interactions did not demonstrate significant causalities, notably in Figure 25 C, D, E, H and I between anomalies SPEI-1 and SGWI-1, SPEI-1 and NDWI, SPEI-1 and NIRv, SPEI-3 and NDWI as well as between SPEI-3 and NIRv respectively. Both causation interactions between SPEI-1 and SGWI-1, as well as SPEI-3 and SGWI-1, showed the convergence of ϱ however, the ϱ score at the maximum library size was lower than the linear correlation indicating spurious causation.

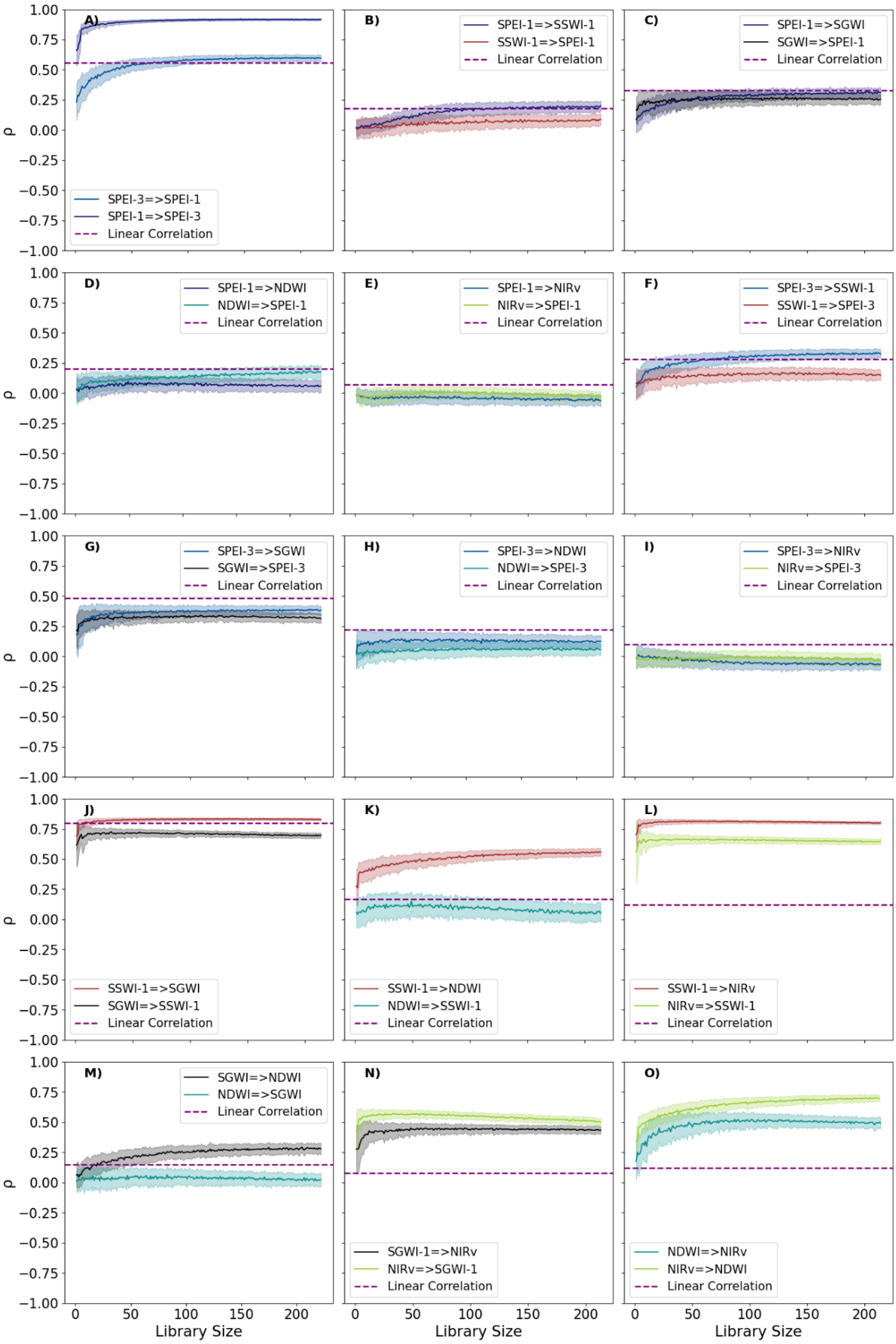


Figure 25: Causal analysis results between the drought and vegetation anomalies.

To summarize the drought propagation from the causation results above, the most influential drought anomaly for each drought type was determined as the one with the highest ϱ score among the analyzed interactions. Figure 26 demonstrates the summary of drought propagation in the study area by displaying the direction of the dominant causal interactions. The demonstrated summary of the dominant drought causations indicates that drought propagation in the Witharen region is initiated by a monthly meteorological drought anomaly which propagates into a 3-month anomaly (SPEI-3). Despite a weak but significant causal influence SPEI-3 propagates to SSWI-1. Parallelly, SSWI-1 propagates to SGWI-1 and NIRv anomalies, with a relatively strong causation influence in both (ϱ >0.8). Finally, NIRv anomalies propagate into NDWI anomalies.

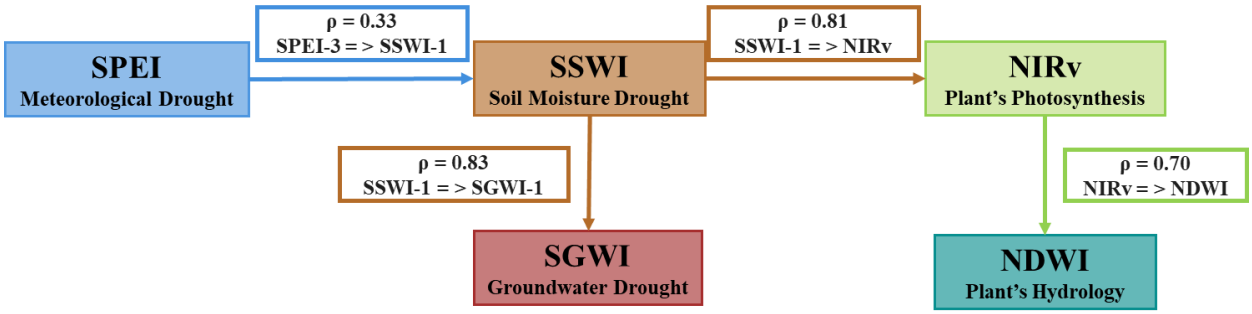


Figure 26: Summary of drought causalities propagation.

4.4.2 Soil Moisture Causal Interactions.

This sub-section presents the results of assessing the drought causal interactions between Surface Soil Moisture anomaly (SSWI-1u) and other drought anomalies. In general, Figure 27 indicates that CCM was able to capture two significant bidirectional causations between SSWI-1u and vegetation anomalies with SSWI-1u being the dominant causal index. Specifically, SSWI-1u=>NIRv and SSWI-1u=>NDWI scored a cross-mapping skill of 0.71 and 0.49, the highest cross-mapping skill in their respective causal interactions. Similar to the causal results of RZSM anomalies, NIRv had the strongest causal relationship with SSM anomalies than NDWI. Additionally, unidirectional causal relationships were detected between SSWI-1u and SGWI-1 as well as SPEI-3. A similar interaction type was also observed with RZSM anomalies suggesting a similar response shown by surface and root zone soil moisture.

Divergence was however noted in the strength of the causations between SSWI-1 and SSWI-1u causation interactions. Particularly, vegetation and groundwater anomalies scored higher cross-mapping skills with SSWI-1 while SPEI-3 demonstrated the highest cross-mapping skill values with SSWI-1u. The causal interaction of SSWI-1u on NDWI and NIRv, scored a ϱ value of 0.49 and 0.71 while SSWI-1 scored 0.81 and 0.53 respectively. A similar trend was observed in interactions with SGWI-1, SSWI-1 scored ϱ value of 0.86 while SSWI-1u scored 0.75. In the case of SPEI-3 however, SSWI-1u scored 0.4, a surpassing ϱ value of 0.33 scored by SSWI-1. This indicates that the surface soil moisture time series is more coupled with the surface dynamics while the root zone soil moisture time series is more coupled with subsurface hydrological dynamics.

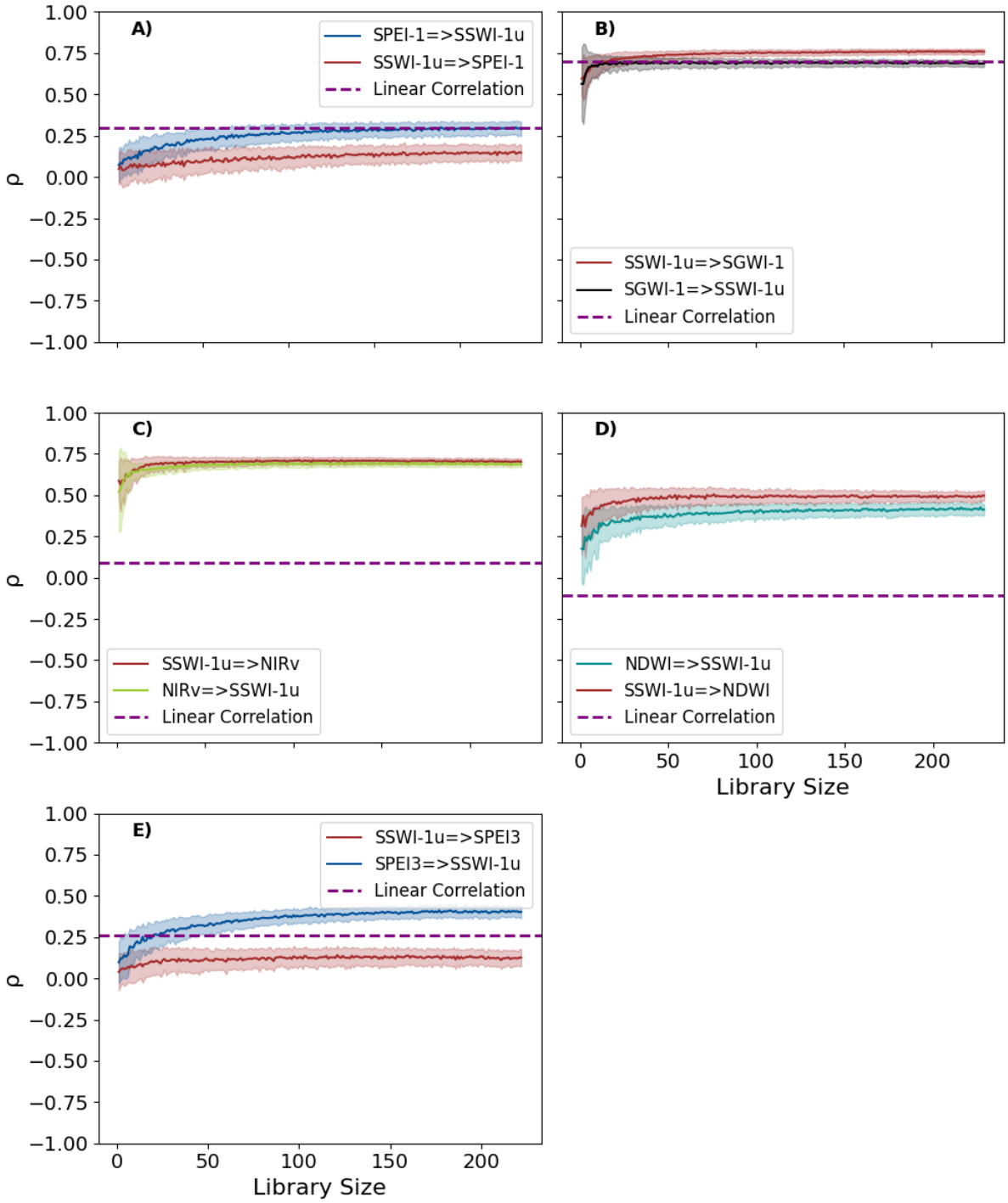


Figure 27: Causal analysis results between the drought and vegetation anomalies. Surface Soil Moisture anomaly was used to compute SSWI-1 instead of Root Zone Soil Moisture time series.

5 DISCUSSIONS

Understanding local drought development processes is crucial for monitoring and forecasting, thereby improving community resilience to drought impacts (X. Zhang et al., 2022). Studies have consistently identified key factors for drought propagation analysis, including delayed response, accumulated impacts, and the demonstrated capability of Root Zone Soil Moisture (RZSM) to infer drought propagation to vegetation anomalies (Dai et al., 2022; Ho et al., 2021; Wu et al., 2018). Despite the understanding of these key features, issues such as the limited spatially detailed and continuous RZSM datasets have restricted the investigations on the drought propagation to vegetation anomalies (Zhuang et al., 2020). Moreover, linear correlation methods are still limited in detecting causal directions (Grillakis et al., 2021).

To address these two issues, this analysis discusses i) the justifications for using 1KmSSM, ii) evaluating the two RZSM estimation methods iii) the general trend of the local computed drought anomalies iv) the linear association between vegetation and drought anomalies, and v) results using Convergent Cross Mapping (CCM) and RZSM to infer local drought and vegetation anomalies relationships and finally vi) The relevance of RZSM and SSM anomalies in drought monitoring.

5.1 Justifications for Using 1KmSSM

The most used standard validation measure for grid soil moisture datasets is Root Mean Square Error (RMSE) to quantify the deviation from the in-situ soil moisture. The benchmark RMSE score of 0.04 m^3/m^3 imposed by global scale soil moisture datasets such as SMAP and SMOS has also been implemented in most soil moisture prediction studies at a local scale (Entekhabi et al., 2010; Kerr et al., 2010). At the Twente soil moisture monitoring stations, the average RMSE score by 1KmSSM grid dataset among the 15 evaluated soil moisture monitoring stations was 0.1 m^3/m^3. Because of the low accuracy, only stations whose 1KmSSM time series was able to score a linear correlation coefficient >0.6 were used for the calibration of RZSM estimation methods. Nevertheless, two stations out of the selected 9 stations had an RMSE above 0.149 m^3/m^3 highlighting further deviation of 1KmSSM from in-situ soil moisture measurements.

It is worth noting, however, that 1KmSSM is a recently published dataset, meaning a limited study on the bias and applicability on a local scale has been conducted. Nevertheless, despite the limitations, analysis by Han et al. (2023), supported the use of 1KmSSM in this study. A detailed comparison with other global long-term soil moisture datasets such as SMAP and ESACC106.1 revealed that 1KmSSM consistently outperformed other datasets in terms of residual measurements at Twente monitoring stations. Additionally, 1KmSSM has proven effective in capturing the spatial variability of soil moisture in temperate regions and was also able to detect the onset and offset of soil moisture anomalies, crucial for soil moisture drought monitoring.

5.2 The Most Effective Method for RZSM Estimation

The saturation of the root zone is the key determinant of soil water availability for plant absorption (Gautam, 2018). Accurate monitoring of the Root Zone Soil Moisture (RZSM) is therefore crucial for the prediction of plants' response to various water availability conditions. The development of remotely sensed Surface Soil Moisture measurements (SSM) has provided an opportunity to reduce costs associated with the installation of soil moisture monitoring stations while improving spatial coverage (Baldwin et al., 2019; Gautam, 2018). However, most spatial detailed datasets monitor soil moisture just up to 5cm deep (Wakigari

& Leconte, 2022). In this objective, the performance of two methods to infer RZSM from the SSM time series is compared against the in-situ measurements at Twente soil moisture monitoring stations to determine the most optimal method for RZSM estimation from the grid SSM dataset.

The results of this objective indicated that the Exponential Low Pass Filter (ELPF) method is more suitable for RZSM estimations from 1KmSSM grid data in the Twente region compared to the Cumulative Distribution Function (CDF) matching method. ELPF consistently produced the most desirable accuracy scores in the seven soil moisture stations used for calibration and accuracy evaluation, (Figure 15). These findings are consistent with observations made by Tian et al. (2020) and X. Gao et al. (2019). For instance, Tian et al. (2020) found that ELPF outperformed both CDF matching and Artificial Neural Network (ANN) by best capturing the subsurface soil moisture time series accurately in a watershed-dominated grassland and scattered woodland. In this study, a less accurate SSM time series input (1KmSSM) (figure 3) is the likely reason for the poor performance by CDF matching whose accuracy is mostly affected by the linear coupling between SSM and RZSM (Z. Xu et al., 2022).

Additionally, the analysis examined the generalizability of the ELPF method, finding that the calibrated ELPF method can be generalized and transferred to other uncalibrated stations while maintaining reasonable RZSM estimation results. Specifically, no significant change in the accuracy metrics score was observed after adopting the depth-specific median optimal T values for all stations as the generalized local-specific T parameter on both calibrated and uncalibrated stations. Similar results were reported by Tian et al. (2020) by using the median T value across a homogenous climatic region. Considering also the reported CDF matching requirements for the calibrated method parameters at each pixel (X. Gao et al., 2017; M. Li et al., 2023), these findings underscore the superiority of ELPF over CDF matching and its applicability in the current study which requires the transferable RZSM estimation model.

5.3 The Spatial-Temporal Pattern of the Drought Indices

Building on the improved aspects of drought monitoring, run theory has been used for quantitative drought identification, monitoring, and characterization (L. Wang et al., 2020). As the means to obtain prior knowledge of drought events in the study area, this section discusses the results of the run theory application on the three drought anomalies time series spanning 2000-2020. The time series analysis of the two properties, drought frequency, and drought intensity provides an overview of the drought trends in the study area. Moreover, spatial analysis of the detected drought properties provides an understanding of the influencing factors of drought development.

5.3.1 General Trend of Drought Episodes and Their Properties

The drought characterization results revealed that the study area experienced a higher drought frequency between 2010 and 2020. The last 10 years of the study period were found to have the greatest number of drought episodes as well as the most intense drought episodes. Figure 18 establishes that more than 65% of drought episodes occurred between 2010-2020. This increase in frequency can be attributed to the episodes of extreme precipitation and temperature anomalies in 2011, 2016, 2018, 2019, and 2020 endured in the Netherlands. These findings are consistent with findings by Philip et al. (2020), who explained a climate change-induced increase in drought frequency from 1950 to 2018. Furthermore, the projected climate change trends discussed by Van Den Hurk et al. (2014) and Christian et al. (2023) support the observed higher drought frequency in the past decade.

Additionally, the most intense meteorological, soil moisture, and groundwater drought episodes were detected in 2011, 2018, and 2019 respectively. Figure 17 shows that the drought episode in 2011 featured a highly negative SPEI-1 anomaly that only lasted for three months. Guardamino et al. (2021) similarly identified the 2011 drought as a brief and extremely dry episode detected by the meteorological drought index at 1, 3, and 6-month scales only. This short-lived SPEI negative anomaly followed by rapid recovery to a positive anomaly explains the less severe SSWI and SGWI negative anomaly. In contrast, droughts of 2018 and 2019 experienced prolonged episodes of SPEI negative anomalies that propagated to severe SSWI and SGWI negative anomalies. This observation is supported by the reported climatic records of extreme temperature and evapotranspiration rates during the summer of 2018 intensified the propagation of precipitation deficiency to groundwater and soil moisture anomalies (Philip et al., 2020).

5.3.2 Spatial Patterns of the Drought Occurrence and Development

The analysis of drought indices and the influencing factors indicated that the propagation of meteorological drought to groundwater anomalies is mostly influenced by the aquifer depth. SGWI time series in areas with shallow aquifers showed more variability compared to locations with deeper aquifers. Also, areas with a shallow aquifer experienced the greatest number of groundwater drought episodes which were rather shorter suggesting a quick feedback interaction between unsaturated and saturated zones. Moreover, a higher positive correlation between SPEI and SGWI was observed in locations with a shallow aquifer compared to the deeper ones suggesting a strong link. Similar observations were made by Brauer et al. (2018) in the Hupsel Brook catchment in which more coupled vadose and saturated zone water dynamics were observed in shallow aquifers suggesting the restoration of the evapotranspiration impacts in these areas by capillarity action from the water table as well as quick groundwater recharge.

However, the current analysis detected no significant association between SWI correlations with both SPEI and SGWI or the intensity of SSWI drought and groundwater depth. This observation contradicts the findings above which associated depth with the interaction between saturated and unsaturated water dynamics. Moreover, it also contradicts the conclusion that the capillarity movement of groundwater from the shallow water table to the root zone is significant, especially during soil moisture drought episodes (Mu et al., 2021). Additionally, soil texture and land cover properties influence soil water-holding capacity and topsoil permeability respectively. However, in this study, limitations associated with accurately assigning land cover classes to 1Km-pixel and a uniform sand texture across the study area couldn't explain the variation in drought correlations.

5.4 Vegetation Condition Anomalies Under Drought Conditions

Vegetation is the critical component of terrestrial ecosystems, forming a natural link for the water-carbon-energy exchange between the atmosphere and soil. By influencing surface albedo and through processes such as photosynthesis and evapotranspiration, vegetation influences water and energy circulation (Chi et al., 2020; Kong et al., 2020). Therefore, understanding the dynamics of vegetation conditions in various water availability conditions is crucial for effective water resources and ecosystem health management. In this sub-section, the time series of two vegetation condition indices were analyzed to reflect on the dynamics of the distinct vegetation properties during drought conditions.

The linear correlation analysis indicated an overall weak linear association was consistent between NDWI and all three drought indices (SPEI-1, SSWI-1, and SGWI-1), whereas NIRv exhibited a similar relationship with soil moisture anomalies. Consistently, both vegetation anomalies showcased the strongest linear correlation with soil moisture anomalies followed by precipitation and then groundwater anomalies.

Moderate correlations were observed during the spring and summer seasons, which are the periods of active vegetation growth and the strongest linear link to drought anomalies.

Furthermore, time series analysis revealed a less intuitive interpretation of the association between drought and vegetation anomalies. A periodic coupling for instance during most of the summer periods indicated a seasonal linear response to soil moisture anomalies. However, issues in the time series such as the delayed vegetation responses, and quick recoveries, together with or followed by recurrent negative vegetation anomalies, were also observed during the drought periods of 2003, 2006, 2009, and 2018. Moreover, the median drought and vegetation anomalies time series indicated a decoupling between the drought and vegetation anomalies during the summer drought of 2005. These temporal attributes contribute to the weak linear relationship between vegetation and drought anomalies, highlighting the complexity of drought-vegetation interactions.

The general weak linear link between vegetation and water anomalies is not a novel finding in this study. Studies have highlighted various factors that can account for the asynchronous dynamics of drought and vegetation anomalies. Kong et al. (2020) highlighted the impacts of the vegetation self-regulatory system both during and after a particular drought episode leading to a cumulative and delayed response to water scarcity anomalies. Studies on the vegetation response to water scarcity in humid and temperate climatic zones have indicated a range of time scales required to obtain a peak linear relationship between drought and vegetation anomalies, varied across vegetation types, ranging from 1 to 11 months (Kong et al., 2020; Weng et al., 2023; Zhao et al., 2020).

Additionally, the energy-limited and water-limited soil regimes have been reported to determine the duality of vegetation-to-drought interaction especially in humid temperate areas. (Van Hateren et al., 2021) explained that below the critical soil moisture content, the system is said to be in a water-limited state and vegetation will respond to the changes in moisture content. However, above that, an increase in incoming solar radiation will lead to an increase in vegetation productivity and enhanced evapotranspiration and not soil moisture content. This duality explains the decoupling between vegetation and drought anomalies during the winter season when the system can be said to be energy-limited and coupled during the summer and spring seasons when the system is water-limited.

5.5 Drought Causality Relationships

In line with the expectations of this study, results have indicated a general increase in drought frequency and intensity in the study area in the last decade of the study period. Furthermore, the linear association analysis between drought and vegetation anomalies in the proceeding subsections highlighted the complexity of drought propagation analysis in the Witharen region. Seasonal variability, delayed responses, and periodic recovery of vegetation indices complicated the assessment of the nature of the drought interactions. Furthermore, a moderate linear relationship between drought anomalies was not enough to indicate the direction of interactions. In this sub-section, the results of applying a non-linear model (CCM) to detect drought causation interactions as well as causation directions of the causal influences are discussed in detail.

Despite the assessment complexities identified by the linear association of drought anomalies, CCM analysis successfully captured significant causal relationships that reflect the prior understanding of the local hydrological system. Specifically, the strongest causation interaction following SPEI-1=>SPEI-3 was from SSWI-1 to SGWI-1. This reflects the characteristic shallow aquifer in the study area, with groundwater depth ranging between 0.7 and 3 m, a strong relationship between RZSM dynamics at 0.4 m depth and

groundwater time series is expected. This also explains the significant causal impact of SGWI-1 on vegetation anomalies despite the area being dominated by croplands, shrubs, and grasslands that are expected to have shallow roots. Groundwater buffering influence on the propagation of meteorological drought to vegetation anomalies aligns with the findings by Fawen et al. (2023) in which a Cross Wavelet Transform model was used to identify drought response.

Moreover, CCM analysis also revealed that soil moisture anomaly is the link for the meteorological drought propagation to other types of drought anomalies. The summary of the highest cross-mapping skill scores for each anomaly indicated general drought propagation in the study area begins with accumulated monthly precipitation anomalies, leading to soil moisture anomalies, from which both groundwater and vegetation anomalies are initiated. This supports the argument that the combination of precipitation deficiency and high evaporative demand is the typical initiator of drought events (Wilhite et al., 1985). High temperatures, strong winds, and low humidity coinciding with prolonged periods of precipitation deficit, result in enhanced impacts of drought anomalies on soil moisture, groundwater, and vegetation anomalies. A similar pattern was observed in the Netherlands, where the 2018 summer precipitation anomaly propagated into soil moisture and groundwater drought in the subsequent winter (Philip et al., 2020; Rutger Weijers, 2020).

In addition to the general drought causality propagation, the following discussion explores other drought and vegetation anomaly causality relationships, considering the expected physical relationship and the literature arguments.

a. Soil Moisture and Plants

This study revealed a consistent causal interaction between root zone soil moisture anomalies and both vegetation water content and vegetation productivity anomalies. A significant bidirectional interaction was detected between SSWI-1 and NIRv, while with NDWI a unidirectional interaction was prominent. A consistent relationship between root zone soil moisture anomalies and vegetation conditions echoes findings by Rigden et al. (2020). Similarly, Buitink et al. (2020) in the Twente region, Netherlands, showed that both vegetation water depth and aboveground biomass have a strong but non-linear relationship with soil moisture critical values.

In both interactions, CCM indicated the soil moisture anomalies to vegetation anomalies as the most dominant causal directions suggesting the dependency of vegetation conditions on the root zone soil water potential. Moreover, a weak but significant causal influence of NIRv on SSWI-1 suggests a feedback response of the aboveground biomass to the temporal evolution of root zone soil moisture anomalies. Processes such as rainfall interceptions, stem flow, and throughfalls by the local ecosystem may account for the vegetation influence on the spatiotemporal distribution of soil moisture content (C. Wang et al., 2019).

b. Soil Moisture and Precipitation

In this study, a consistent unidirectional causal interaction was observed between meteorological and soil moisture anomalies, with both 1-month and 3-month precipitation anomalies exerting a significant causal influence on monthly soil moisture anomalies. Conversely, no significant causal impact was observed from SSWI-1 to either SPEI-1 or SPEI-3. This absence of causal influence of soil moisture anomaly on meteorological drought aligns with the observation by Van Der Velde et al. (2021), who noted that both the atmospheric moisture content and temperature of the study area are the product of dry and cold winds blowing from the land mass and moist and warm winds blowing westward from the sea. This suggests that, in a study area, although meteorological drought can propagate to soil moisture anomalies, the desiccation

level of the local soil moisture does not have a significant influence on the dynamics of the local meteorological forcings.

c. Plants and Precipitation

CCM analysis did not find a consistent significant causal relationship between vegetation indices and meteorological drought. Both SPEI-1 and SPEI-3 to NDWI CCM results showed a unidirectional insignificant convergence suggesting a spurious causation interaction. On the other hand, no convergence was observed with both SPEI-1 and SPEI-3 causation interactions with NIRv. These findings contradict the vegetation's bidirectional feedback relationship with temperature and precipitation dynamics involving processes such as changes in latent heat flux, albedo, and surface roughness (Jiang et al., 2020; Kong et al., 2020; W. Zhang et al., 2023).

The dynamic nature of vegetation response to precipitation anomalies could explain the lack of significant convergence. Van Hateren et al. (2021) explained vegetation responses to precipitation anomalies are significantly weakened during energy-limited regimes in temperate and humid areas. A decline in photosynthesis during the energy-limited regimes in humid and temperate regions is likely to render the vegetation conditions irresponsive to the precipitation anomalies.

Additionally, shallow groundwater levels and drainage canals (Figure 28) in the study area are likely to weaken the impacts of precipitation anomalies on the vegetation conditions during the summer season. Verdonschot & Piet (2017) discussed how the 'Room for River' project improved soil water retention during the dry season. However, drainage canals are well distributed in the study area posing a challenge in the comparison of their impact on the coupling of drought indices at a 1 Km resolution. These findings suggest that despite a significant positive seasonal coupling between meteorological drought and vegetation anomalies (figure 22), at a monthly scale and 1 km resolution, vegetation anomalies do not exhibit a causal interaction with meteorological drought anomalies.

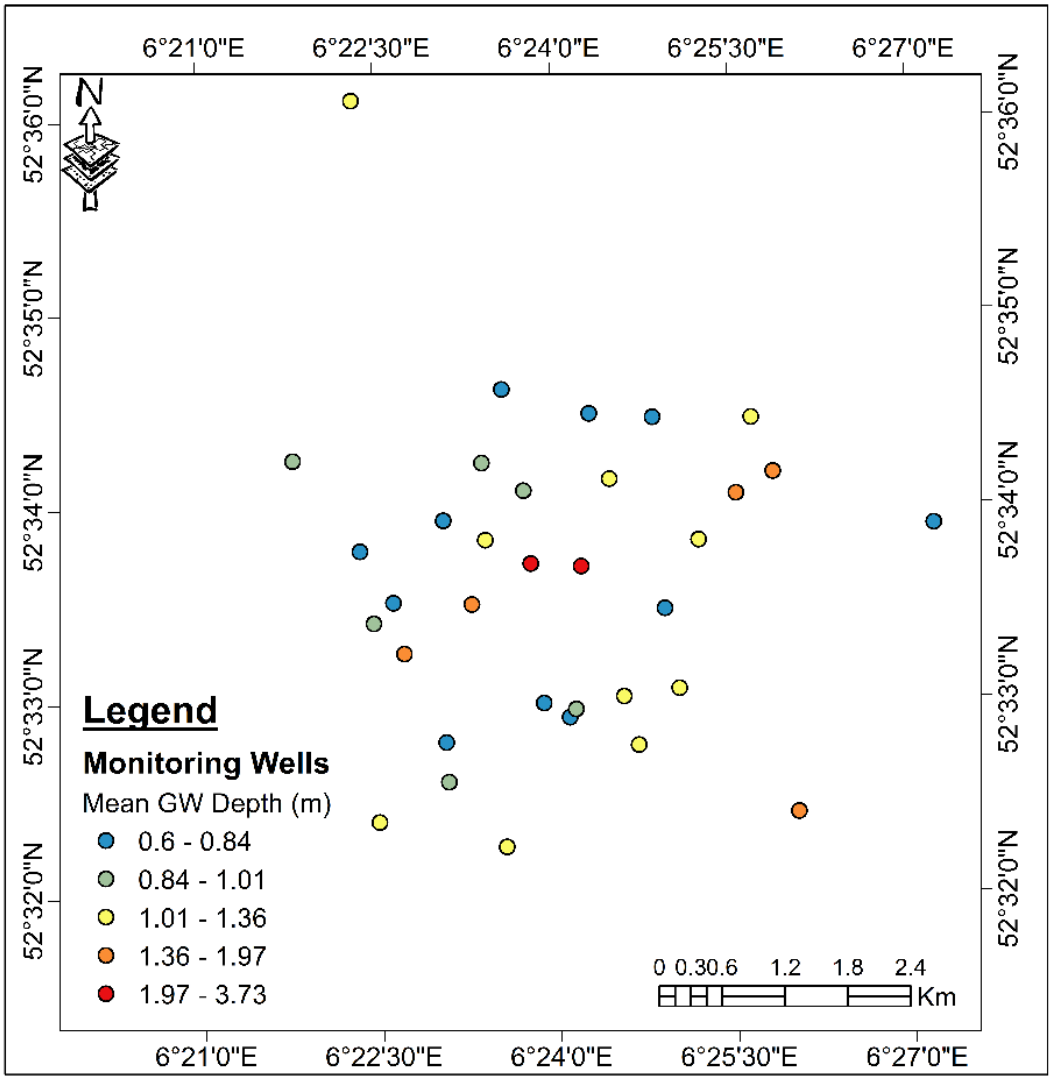


Figure 28: Average groundwater depth for the Witharen monitoring network)

d. Plants and GW

CCM results also revealed that both NDWI and NIRv time series can cross-map SGWI-1 at a significant convergence indicating significant groundwater causal impacts on vegetation conditions. This is consistent with the study area's shallow aquifer properties. Higher cross-map skill was scored by the SGWI-1=>NIRv than SGWI-1=>NDWI causation direction suggesting a stronger response by NIRv to groundwater anomalies compared to NDWI. In both interactions, the causation direction was, however, unidirectional suggesting that although groundwater influences vegetation condition, at a 1 Km and monthly scaled analysis, no vegetation to groundwater anomaly causations was detected.

These findings align with the interpretation of the study area properties by the Advies Commissie Schade Grondwater (ACSG) in the Netherlands. According to ACSG (2023), groundwater storage influences vegetation conditions through capillarity water movement to the soil's root zone, especially in periods of dry precipitation anomalies. This buffering impact is much more pronounced in shallow aquifers (Rutger Weijers, 2020), for instance, the Witharen region.

Moreover, the interaction can be bidirectional considering how the evapotranspiration and interception by vegetation influence groundwater levels and distribution (Q. Zhang et al., 2022). However, the lack of significant convergence in both NDWI and NIRv to SGWI-1 CCM causal analysis in the current analysis suggests the presence of other dominant factors influencing the groundwater dynamics.

e. Precipitation and GW

Groundwater drought is typically thought to develop from precipitation anomalies through soil moisture drought but with a delayed and complicated response (Leelaruban et al., 2017; Petersen-Perlman et al., 2022). In the current study, a medium to high linear correlation was observed between meteorological drought anomalies and SGWI-1. However, no consistent significant causation was detected by CCM causal analysis. Despite the convergence, the cross-mapping skill at a maximum library size was lower than the linear correlation suggesting a spurious causation. Nevertheless, SPEI-3=>SGWI-1 had a higher cross-mapping skill at the maximum library size compared to SPEI-1=>SGWI-1 suggesting increased impacts with prolonged precipitation anomaly.

Despite the shallow aquifer, the commercial groundwater exploration activities in Vitens groundwater extraction activities in Witharen might be the primary influencing factor for SGWI-1 anomaly hence weak causation by SPEI. At a certain rate, groundwater extraction can induce groundwater depression leading to deep groundwater flow. For instance, during the drought period of 2018, water production from this area was reported to increase by 40% to accommodate demands leading to the inflated risk of salinization (Van Wijnen & Schotting, 2019). Similarly, linear coupling between SGWI-1 and SPEI-1 was higher in shallow wells, away from the extraction area than in deeper wells which are close to the extraction zone (section 4.2). However, the impact of groundwater extraction on the decoupling of SPEI and SGWI in the Witharen region remains to be tested and quantified.

f. NIRv and NDWI

The CCM analysis in this study found a moderate linear correlation and a consistent bi-direction significant causation between NDWI and NIRv indicating a feedback-like mechanism. NIRv=>NDWI causation direction scored the highest cross-map skill compared to the opposite causation direction. Similarly, with causal analysis involving SSWI-1 and SGWI-1 anomalies, vegetation productivity had shown a stronger causal response than vegetation water content. Buitink et al. (2020) indicated similar results when comparing the response of the optical vegetation productivity and vegetation water content to the soil moisture drought using daily time series datasets. Buitink's analysis conducted in Overijssel revealed that at the beginning of the soil moisture drought, vegetation water content was increasing while NIRv kept up with the soil moisture anomalies. Furthermore, vegetation water content took a longer time to recover compared to vegetation productivity.

Vegetation expresses a complicated response mechanism to drought stress that can be simplified into a simple model that includes processes such as stomata closure and shading leaves that preserve water but impact productivity (Farooq et al., 2009). The mechanism is even more complicated considering the stages of the response and variation among species (Farooq et al., 2009; Yang et al., 2021). The data-driven CCM causal analysis results in this study indicated that vegetation productivity experiences a stronger causal interaction with drought anomalies compared to vegetation water content suggesting a stronger link between drought and vegetation productivity, whereas vegetation regulatory influences can be linked to the weakens the vegetation water content response to drought anomalies.

5.6 The Relevance of RZSM and SSM Anomalies for Detecting Drought Causalities.

In line with the hypothesis that RZSM anomaly has the most causal influence on the vegetation anomaly than SSM anomaly, in this analysis, the implications of both SSM and RZSM in capturing various drought relations in the study area are evaluated. Depth distribution of plants' roots of most crops (Fan et al., 2016; Foxx et al., 1984) suggests that RZSM can be associated the most with the water content available for plant growth (Grillakis et al., 2021). On the other hand, sensible and latent heat flux at the surface level has been more associated with the moisture dynamics on the soil surface than on the root zone (Qing et al., 2023; Taylor et al., 2012). The relevance of the two-time series in the detected drought interactions is analyzed in this sub-section.

Similarity in the detected relationships was observed with drought association analysis using either SSM or RZSM anomalies. A consistent similarity in the correlation score signs with other drought anomalies highlighted the similarity in linear relationships. Moreover, CCM results found that both SSM and RZSM anomalies had a unidirectional causal interaction with both groundwater and precipitation anomalies while with NIRv a bidirectional causal interaction was prominent. This suggests the uniformity in SSM and RZSM causal interactions with other anomalies.

However, the variation in relations was observed with the magnitude of the correlation and cross-mapping skills. SSM anomalies demonstrated a stronger association with SPEI-3, whereas RZSM showed more coupling with groundwater and vegetation anomalies. These findings align with the observations that SSM has been on the surface is more sensitive to precipitation and evapotranspiration processes compared to RZSM which is more related to sub-surface activities such as infiltration, capillarity of groundwater, and the soil water absorption by plants' roots.

Furthermore, a stronger association with SGWI-1 reflects on the proximity in the sub-surface environment, where the root zone soil moisture (typically at 40cm depth) influences the dynamic of groundwater level (average range 0.7 to 3 m in the study area). On the other hand, a stronger coupling between SPEI-3 and SSM than with RZSM anomalies was expected as both precipitation and evapotranspiration impact the earth's surface before developing into sub-surface dynamics through processes such as capillarity and infiltration.

In light of these findings, it is evident that both SSM and RZSM play integral roles in the detection of drought causalities, each reflecting different aspects of the soil moisture continuum and its interaction with other drought variables.

6 LIMITATIONS OF THE RESEARCH

This study was able to assess the causality interactions between drought and vegetation condition anomalies. By analyzing the predictability of the causal anomaly based on the signals on the time series of the affected anomaly significant causal interactions were detected. Although most of the detected causations aligned with the general knowledge of the study area, it is crucial to highlight that, the detected causal interactions are purely data-driven. A comprehensive analysis of the underlying physical process will be crucial to support these findings. Additionally, this study didn't account for the seasonality and the long-term trend of the causal interactions.

This study is fundamentally dependent on the availability of both in-situ measurements and accurate remote sensing estimates of drought-related variables. To replicate this study in a different geographical context, there are specific data requirements. Firstly, at least three years' worth of in-situ measurements of RZSM for the calibration of both the CDF matching and ELPF methods, as outlined by X. Gao et al. (2017). Secondly, a comprehensive and continuous groundwater monitoring network is required, robust enough to generate at least 100-time series match-ups with other drought variables. The data size was identified by Kurths & Herzel (1987) and Smith (1988) as the minimum data requirement to establish causation using CCM.

Moreover, this study has endured various sources of uncertainty that may limit the application of the highlighted findings. Der Kiureghian & Ditlevsen (2008) classified uncertainty into two main classes intrinsic randomness and systematic errors. Intrinsic randomness, also called aleatoric uncertainty, is inherited from the dataset while systematic also called epistemic uncertainty can be reduced as it arises from a lack of enough knowledge about the system or model uncertainty. In this section, some of the main sources of uncertainty observed in this study are presented as well as the potential alternatives with less emphasis on classification.

6.1 Spatial Scaling

The main source of systematic uncertainties in this study arose from the low-resolution spatial continuous datasets used for the monitoring of the drought indices. Using course resolution datasets, the retrieved time series of soil moisture drought, and vegetation anomalies provided only the generalized compounding overview of the variables of interest that can be misleading in some areas within the 1 Km surface grid.

Furthermore, the heterogeneity of the spatially governing factors at the grid scale compounds the uncertainties of the generalized physical parameter (Lee et al., 2017). For instance, 1KmSSM time series retrievals from stations surrounded by highly mixed land cover within a 1km grid scale demonstrated a relatively high mean residual distance. For example, ITC_SM01 containing two Dinkel River tributaries together with grasslands and a tarmac road scored an RMSE of 0.14 m^3/m^3, similar observations were noted with ITC_SM05 and 14. This limitation is hence more relevant to the replication of the method in areas with a high degree of land cover heterogeneity within a 1 km grid scale.

Moreover, the coupling between in-situ measurements of groundwater and grid values of soil moisture and optical indices aggregates the scaling dilemma. The in-situ time series presents the temporal records of the parameter of interest at a specific point in space whereas, the grid datasets provide a grid-scale generalized

overview time series. Consequently, this spatial discrepancy may impact the coupling between the time series. Likely, the uncertainty in the spatial representation can be propagated to the causal analysis hence impacting the detected causation results.

The most viable alternatives would involve the use of in-situ measurements of the drought parameters. However, this solution is constrained by the limited availability of in-situ measurements as well as the uneven distribution of the available monitoring stations. Another solution will be the use of high spatial resolution raster datasets such as X. Li (2023) and Shorachi et al. (2022), however, the temporal coverage limits the application of such datasets in drought monitoring. A well-documented option involves upscaling or downscaling using spatially regressed models (Cai et al., 2022; Crow et al., 2012; Tu, 2019; Van Der Velde et al., 2021; Zhuang et al., 2021). However, Lee et al., (2017) discussed the inherent uncertainty from determinant variables which compounds with the increase of target resolution.

6.2 Temporal Dynamics

Temporal attributes of the datasets also introduced uncertainties, impacting various phases of the research. The tradeoff between spatial resolution and temporal coverage of the analysis justified the data choices. However, the computation of drought indices requires the use of a long time series for a more objective interpretation of the anomalies, preferably 30 years (AghaKouchak, Farahmand, et al., 2015; Su et al., 2017). Data choices in this study limited the analysis to 20 years meant that potential errors in the presentation of climatological mean conditions could be expected. Similarly, the monthly temporal scale risks the possibility of generalizing important drought information such as the duration, response time, and cumulative effects of drought.

The issues caused by the temporal coverage of the input datasets are likely to be aggravated with the CCM analysis. Although Ombadi et al. (2020) argued that CCM can establish significant convergence between hydrometeorological variables using at least 100 days long daily time series, various studies have suggested a different minimum data length requirement to establish convergence (Kurths & Herzel, 1987; Smith, 1988). Additionally, Bonotto et al. (2022) found that the strength and significance of CCM convergence between groundwater and streamflow were significantly decreased by the use of monthly time series compared to the results of using daily time series. Moreover, Sugihara et al. (2012) argued that the more enriching time series increases the probability of reconstructing the manifolds and finding the true close neighbors. Therefore, the causal analysis using the 20-year, monthly datasets may influence the significance of the convergence, and strength of the cross-map skill hence the direction of drought causation may also be impacted.

6.3 Other Sources of Uncertainty

Other sources of uncertainty were introduced by the methods of choice.

- First, both the ELPF and CDF matching method utilizes the coupling between surface and subsurface moisture dynamics to infer RZSM from SSM. Rendering the RZSM estimation is vulnerable to natural or human-induced decoupling as well as changes in the complexity of the SSM-RZSM (AghaKouchak, Feldman, et al., 2015; Rutger Weijers, 2020).
- Second, drought indices particularly SGWI and SSWI anomalies were both computed at a monthly scale to infer drought episodes in the time series. However, Guardamino et al. (2021) have indicated the application of the multiscaled monthly anomalies to provide more details in drought monitoring

and detecting the most propagated drought events. W. Zhang et al. (2023) indicated how various scales of drought indices impact the coupling which can be translated into the causal interactions.

- Additionally, SPEI, SSWI, and SGWI were used for the monitoring of meteorological, soil moisture, and groundwater drought respectively. However, drought impacts a range of variables that are not limited to precipitation, root zone soil moisture, and groundwater only allowing different approaches to drought quantification.
- Moreover, CCM assumes a deterministic system to detect significant causality. Ombadi et al. (2020) highlighted the susceptibility of CCM to noises and stochastic interactions, suggesting its superiority in less chaotic time series. In this study, considering issues with the accuracy of the input datasets, the uncertainty estimation of the CCM results remains the decisive task for the future.
- Furthermore, this study is also vulnerable to errors associated with the optical vegetation indices used in the computation of vegetation anomalies. Both Zeng et al. (2022) and B. C. Gao (1996) highlighted the potential unquantified impacts of soil water absorption, clouds, and snow cover on the NDWI time series furthermore NIRv time series also contains NDVI signals which were also used for the creation of 1KmSSM suggesting the unquantified impacts on the causal signals with SSWI-1 time series.

Additionally, it is worth noting that in-situ Twente soil moisture measurements used as the objective measure of the RZSM estimation accuracy, are themselves subject to uncertainty. The impacts of the land cover and local drainage differences associated with the measurement setup of the soil moisture stations at the edge of the agricultural fields as addressed during field campaigns of 2016 affect the long-term objectivity of the measurements. Furthermore, general issues such as probe calibration inconsistency and the varied impacts of the soil depth and root distribution have been discussed to impact the objective comparability with in-situ soil moisture measurements (M. Li et al., 2023; Tian & Zhang, 2023; Van Der Velde et al., 2022).

In general, the discussed findings of this study underscored the potential of detecting drought-causal interactions in a local catchment. Emphasis on the use of the non-linear Convergent Cross-Mapping (CCM) method and the applicability of both surface and root zone soil moisture anomalies to capture various dimensions of the process. The potential to retrieve RZSM time series from grid SSM datasets was also evaluated as well as the relevance of using both SSM and RZSM in the drought propagation analysis. Despite the physically explainable results, a conservative approach to the findings should be preferred considering the discussed limitations.

7 CONCLUSION AND RECOMMENDATIONS.

7.1 Conclusion

This study has shed light on the potential of unraveling the general interaction of drought anomalies through the detection of causal relationships between drought indices and vegetation condition anomalies within the local context, for this case the Witharen groundwater monitoring network. Initially, the study underscored the complications associated with the interpretation of the linear correlation analysis of drought interactions, emphasizing the need for a non-linear drought analysis to understand interactions among drought and vegetation anomalies in the study area.

The key finding of this study is the demonstrated ability to detect reasonable causal relationships between drought anomalies and vegetation conditions. The detected relationships reflected the general knowledge of the physical and hydrological processes of the study area despite the highlighted complexities of linear correlation interpretations. A notable weak influence of precipitation can mostly be attributed to commercial groundwater extractions and water management activities.

Additionally, this study successfully calibrated a localized transferable method that can infer Root Zone Soil Moisture (RZSM) time series using grid Surface Soil Moisture (SSM) with reasonable accuracy. However, less accurate results were observed in pixels with less precise grid SSM estimates. Moreover, the application of RZSM in the analysis of soil moisture drought interaction demonstrated superiority in detecting causal impacts on vegetation and groundwater anomalies, whereas, for meteorological drought propagation, SSM demonstrated more superiority.

The highlighted findings of this study can be used to enhance the understanding of the general drought interactions in the Witharen region, consequently improving insights into drought forecasting and preparation. Knowledge of the most influential process for local water resources and vegetation conditions will help agricultural and drought authorities plan for rational drought mitigation and adaptation plans. For instance, in the Witharen region, soil moisture anomaly was at the center of the propagation of other drought anomalies from the meteorological anomaly.

Moreover, the approach used in this study can be replicated in other regions where the analysis of drought interactions with vegetation anomalies will be required, especially in areas where linear correlation analysis failed to provide conclusive results of the drought associations or where the knowledge for causal direction is of more importance than just linear association. This approach can also be beneficial for studies that require RZSM time series but are limited due to non-uniform soil moisture monitoring stations. In such cases, the study can serve as a blueprint for the research design. However, issues such as the availability of the in-situ measurements of soil moisture and groundwater, temporal coverage as well as chaotic time series may limit both the calibration of the two described RZSM estimation methods and the CCM ability to infer causation of drought types.

To sum up, this study not only contributes to the understanding of the interaction between drought types but also proposes an effective methodology for future research in the field of drought monitoring. The findings of this study have shown that nonlinear analysis and the use of multiple drought indices can provide a better understanding of drought dynamics and their effects on vegetation conditions. This research, therefore, holds significant implications for drought forecasting, preparation, and management strategies.

7.2 Recommendations

Building from the limitations of this study, any future research on the subject could address the following issues to strengthen drought causation analysis.

I. Utilizing a more precise surface soil moisture dataset could ensure a more reliable estimate of soil moisture anomalies. Areas with a high network of soil moisture monitoring stations will be ideal for this considering most of the publicly available global grid soil moisture datasets have a very coarse spatial resolution.

II. Utilize a finer spatial resolution analysis that could strengthen the causal results by reducing uncertainties introduced by a coarse resolution dataset. Similarly, to achieve this a long-term record of a network of soil moisture monitoring stations would be ideal.

III. For RZSM estimation, use a less coupling-dependent method to reduce the impacts of surface-to-root zone soil moisture decoupling on the resultant time series. For instance, machine learning models such as Random Forest regression have demonstrated a promising capacity to estimate RZSM time series (Carranza et al., 2021).

IV. Assessing the robustness of the detected causalities by testing the impacts of temporal resolution on the detected causal interactions. Using datasets utilized in this study, one approach could focus on soil moisture, groundwater, and precipitation anomalies which are available daily from 2006 to 2020.

V. Validation of the detected causal interactions using physically informed models. For instance, MODFLOW and GSFLOW have been implemented to model the impacts of precipitation, soil moisture, and vegetation on groundwater levels (Das et al., 2023; Martínez-De La Torre & Miguez-Macho, 2019).

VI. Performance comparison: The drought causation results from Convergent Cross-Mapping (CCM) used in this study could be compared against results from other less deterministic drought causation methods such as machine learning, wavelet transform, probabilistic, and piecewise correlation approaches (Chi et al., 2020; Q. Li et al., 2019; Y. Xu et al., 2021).

VII. A sensitivity analysis of the causality approach to factors such as groundwater depth, land cover properties, and climate scenarios can be tested by replicating the approach in various locations with varied properties.

VIII. Utilize methods such as CDF matching to first undertake bias treatment on the grid SSM dataset at a local level to improve soil moisture monitoring (L. Wang & Chen, 2014).

8 ETHICAL CONSIDERATIONS

This study assesses the causal relationship between drought and vegetation anomalies in Overijssel province, Netherlands. The study did not generate, use, or collect personal information. The datasets utilized in this study are publicly available. Published in-situ and grid soil moisture measurements were properly cited and referenced in the reference list.

Furthermore, for open science, the Data Management Plan (DMP) is provided to highlight the details of the storage and sharing of the datasets used in this study.

9 REFERENCES

ACSG. (2023). *Het invloedsgebied van grondwateronttrekkingen voor droogteschade*.

Actuell Hoogetebestand Nederland. (2015). *Handleiding "AHN downloaden van PDOK."* www.ahn.nl

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

AghaKouchak, A., 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. In *Reviews of Geophysics* (Vol. 53, Issue 2, pp. 452–480). Blackwell Publishing Ltd. https://doi.org/10.1002/2014RG000456

AghaKouchak, A., Feldman, D., Hoerling, M., Huxman, T., & Lund, J. (2015). Water and climate: Recognize anthropogenic drought. *Nature*, *524*(7566), 409–411. https://doi.org/10.1038/524409A

Albergel, C., ¨ Udiger, C. R., Pellarin, T., Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Piguet, B., & Martin, E. (2008). Hydrology and Earth System Sciences From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations. *Hydrol. Earth Syst. Sci*, *12*, 1323–1337. www.hydrol-earth-syst-sci.net/12/1323/2008/

Baldwin, D., Manfreda, S., Lin, H., & Smithwick, E. A. H. (2019). Estimating root zone soil moisture across the Eastern United States with passive microwave satellite data and a simple hydrologic model. *Remote Sensing*, *11*(17). https://doi.org/10.3390/rs11172013

Bonotto, G., Peterson, T. J., Fowler, K., & Western, A. W. (2022). Identifying Causal Interactions Between Groundwater and Streamflow Using Convergent Cross-Mapping. *Water Resources Research*, *58*(8), e2021WR030231. https://doi.org/10.1029/2021WR030231

Brakkee, E., Van Huijgevoort, M., & Bartholomeus, R. P. (2021). Spatiotemporal development of the 2018-2019 groundwater drought in the Netherlands: a data-based approach. *Hydrology and Earth System Sciences*. https://doi.org/10.5194/hess-2021-64

Brakkee, E., Van Huijgevoort, M. H. J., & Bartholomeus, R. P. (2022). Improved understanding of regional groundwater drought development through time series modelling: the drought in the Netherlands. *Hydrol. Earth Syst. Sci*, *26*, 551–569. https://doi.org/10.5194/hess-26-551-2022

Brauer, C. C., van der Velde, Y., Teuling, A. J., & Uijlenhoet, R. (2018). The Hupsel Brook Catchment: Insights from Five Decades of Lowland Observations. *Vadose Zone Journal*, *17*(1), 1–8. https://doi.org/10.2136/VZJ2018.03.0056

Buitink, J., Swank, A. M., Van Der Ploeg, M., Smith, N. E., Benninga, H. J. F., Van Der Bolt, F., Carranza, C. D. U., Koren, G., Van Der Velde, R., & Teuling, A. J. (2020). Anatomy of the 2018 agricultural drought in the Netherlands using in situ soil moisture and satellite vegetation indices. *Hydrology and Earth System Sciences*, *24*(12), 6021–6031. https://doi.org/10.5194/HESS-24-6021-2020

Cai, Y., Fan, P., Lang, S., Li, M., Muhammad, Y., & Liu, A. (2022). Downscaling of SMAP Soil Moisture Data by Using a Deep Belief Network. *Remote Sensing*, *14*(22). https://doi.org/10.3390/rs14225681

Carranza, C., Nolet, C., Pezij, M., & van der Ploeg, M. (2021). Root zone soil moisture estimation with Random Forest. *Journal of Hydrology*, *593*. https://doi.org/10.1016/j.jhydrol.2020.125840

Chi, K., Pang, B., Cui, L., Peng, D., Zhu, Z., Zhao, G., & Shi, S. (2020). Modelling the Vegetation Response to Climate Changes in the Yarlung Zangbo River Basin Using Random Forest. *Water 2020, Vol. 12, Page 1433*, *12*(5), 1433. https://doi.org/10.3390/W12051433

Christian, J. I., Martin, E. R., Basara, J. B., Furtado, J. C., Otkin, J. A., L Lowman, L. E., Hunt, E. D., Mishra, V., & Xiao, X. (2023). Global projections of flash drought show increased risk in a warming climate.

COMMUNICATIONS EARTH & ENVIRONMENT, *4*(165). https://doi.org/10.1038/s43247-023-00826-1

Crow, W. T., Berg, A. A., Cosh, M. H., Loew, A., Mohanty, B. P., Panciera, R., De Rosnay, P., Ryu, D., & Walker, J. P. (2012). Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products. *Reviews of Geophysics*, *50*(2). https://doi.org/10.1029/2011RG000372

Dai, M., Huang, S., Huang, Q., Zheng, X., Su, X., Leng, G., Li, Z., Guo, Y., Fang, W., & Liu, Y. (2022). Propagation characteristics and mechanism from meteorological to agricultural drought in various seasons. *Journal of Hydrology*, *610*. https://doi.org/10.1016/J.JHYDROL.2022.127897

Das, B., Singh, S., Thakur, P., & Jain, S. K. (2023). Assessment of future groundwater levels using Visual MODFLOW in the Gomti River basin in India. *Theoretical and Applied Climatology*, *155*(4), 2917–2936. https://doi.org/10.1007/S00704-023-04795-5/FIGURES/16

Delforge, D., De Viron, O., Vanclooster, M., Van Camp, M., & Watlet, A. (2022). Detecting hydrological connectivity using causal inference from time series: synthetic and real karstic case studies. *Hydrol. Earth Syst. Sci*, *26*, 2181–2199. https://doi.org/10.5194/hess-26-2181-2022

Der Kiureghian, A., & Ditlevsen, O. (2008). Aleatory or epistemic? Does it matter? *Structural Safety*, *31*, 105–112. https://doi.org/10.1016/j.strusafe.2008.06.020

Ellen, J. (1992). Growth, yield and composition of four winter cereals. I. Biomass, grain yield and yield formation. *Netherlands Journal of Agricultural Science*, *42*, 153–165.

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

Fan, J., McConkey, B., Wang, H., & Janzen, H. (2016). Root distribution by depth for temperate agricultural crops. *Field Crops Research*, *189*, 68–74. https://doi.org/10.1016/J.FCR.2016.02.013

Farooq, M., Wahid, A., Kobayashi, N., Fujita, D., & Basra, S. M. A. (2009). Plant drought stress: effects, mechanisms and management. *Agron. Sustain. Dev*, *29*, 185–212. https://doi.org/10.1051/agro:2008021

Fawen, L., Manjing, Z., Yong, Z., & Rengui, J. (2023). Influence of irrigation and groundwater on the propagation of meteorological drought to agricultural drought. *Agricultural Water Management*, *277*, 108099. https://doi.org/10.1016/J.AGWAT.2022.108099

Foxx, T. S., Tierney, G. D., & Williams DISCLAIMER, J. M. (1984). Rooting Depths of Plants Relative to Biological and Environmental Factors. *OSTI*.

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

Gao, X., Zhao, X., Brocca, L., Huo, G., Lv, T., & Wu, P. (2017). *Depth scaling of soil moisture content from surface to profile: multi-station testing of observation operators*. https://doi.org/10.5194/hess-2017-292

Gao, X., Zhao, X., Brocca, L., Pan, D., & Wu, P. (2019). Testing of observation operators designed to estimate profile soil moisture from surface measurements. *Hydrological Processes*, *33*(4), 575–584. https://doi.org/10.1002/HYP.13344

Gautam, U. (2018). *Quantifying root zone soil moisture at field scale through downscaling of SMAP Level 4 Soil Moisture product using Sentinel-1*.

Grillakis, M. G., Koutroulis, A. G., Alexakis, D. D., Polykretis, C., & Daliakopoulos, I. N. (2021). Regionalizing Root-Zone Soil Moisture Estimates From ESA CCI Soil Water Index Using Machine

Learning and Information on Soil, Vegetation, and Climate. *Water Resources Research*, *57*(5), e2020WR029249. https://doi.org/10.1029/2020WR029249

Guardamino, L., Su, Z., Zeng, Y., & Rwasoka, D. (2021). *A COMPREHENSIVE COUNTRY-SCALE DROUGHT MONITORING, THE NETHERLANDS*.

Han, Q., Zeng, Y., Zhang, L., Wang, C., Prikaziuk, E., Niu, Z., & Su, B. (2023). Global long term daily 1 km surface soil moisture dataset with physics informed machine learning. *Scientific Data 2023 10:1*, *10*(1), 1–12. https://doi.org/10.1038/s41597-023-02011-7

Han, Z., Huang, S., Huang, Q., Leng, G., Liu, Y., Bai, Q., He, P., Liang, H., & Shi, W. (2021). GRACE-based high-resolution propagation threshold from meteorological to groundwater drought. *Agricultural and Forest Meteorology*, *307*, 108476. https://doi.org/10.1016/J.AGRFORMET.2021.108476

Ho, S., Tian, L., Disse, M., & Tuo, Y. (2021). A new approach to quantify propagation time from meteorological to hydrological drought. *Journal of Hydrology*, *603*, 127056. https://doi.org/10.1016/J.JHYDROL.2021.127056

Ionita, M., Nagavciuc, V., Scholz, P., & Dima, M. (2022). Long-term drought intensification over Europe driven by the weakening trend of the Atlantic Meridional Overturning Circulation. *Journal of Hydrology: Regional Studies*, *42*. https://doi.org/10.1016/J.EJRH.2022.101176

Jiang, P., Ding, W., Yuan, Y., & Ye, W. (2020). Diverse response of vegetation growth to multi-time-scale drought under different soil textures in China's pastoral areas. *Journal of Environmental Management*, *274*, 110992. https://doi.org/10.1016/j.jenvman.2020.110992

Kadir, M., Fehri, R., Souag, D., & Vanclooster, M. (2020). Exploring causes of streamflow alteration in the Medjerda river, Algeria. *Journal of Hydrology: Regional Studies*, *32*, 100750. https://doi.org/10.1016/J.EJRH.2020.100750

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

Kong, D., Miao, C., Wu, J., Zheng, H., & Wu, S. (2020). Time lag of vegetation growth on the Loess Plateau in response to climate factors: Estimation, distribution, and influence. *Science of The Total Environment*, *744*, 140726. https://doi.org/10.1016/J.SCITOTENV.2020.140726

Kurths, J., & Herzel, H. (1987). *An Attractor in a Solar Time Series*. 165–172.

Lee, J. H., Zhao, C., & Kerr, Y. (2017). Stochastic Bias Correction and Uncertainty Estimation of Satellite-Retrieved Soil Moisture Products. *Remote Sensing 2017, Vol. 9, Page 847*, *9*(8), 847. https://doi.org/10.3390/RS9080847

Leelaruban, N., Padmanabhan, G., & Oduor, P. (2017). Examining the Relationship between Drought Indices and Groundwater Levels. *Water 2017, Vol. 9, Page 82*, *9*(2), 82. https://doi.org/10.3390/W9020082

Li, M., Sun, H., & Zhao, R. (2023). A Review of Root Zone Soil Moisture Estimation Methods Based on Remote Sensing. In *Remote Sensing* (Vol. 15, Issue 22). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/rs15225361

Li, Q., He, P., He, Y., Han, X., Zeng, T., Lu, G., & Wang, H. (2019). *Investigation to the relation between meteorological drought and hydrological drought in the upper Shaying River Basin using wavelet analysis*. https://doi.org/10.1016/j.atmosres.2019.104743

Li, X. (2023). *Estimating Daily Soil Moisture at High Spatial Resolution for Drought Monitoring by Fusing Multi-Source Data Based on Random Forest*.

Martens, B., Miralles, D. G., Lievens, H., Van Der Schalie, R., De Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., & Verhoest, N. E. C. (2017). GLEAM v3: satellite-based land evaporation and

root-zone soil moisture. *Geosci. Model Dev*, *10*, 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017

Martínez-De La Torre, A., & Miguez-Macho, G. (2019). Groundwater influence on soil moisture memory and land-atmosphere fluxes in the Iberian Peninsula. *Hydrology and Earth System Sciences*, *23*(12), 4909–4932. https://doi.org/10.5194/HESS-23-4909-2019

Mu, M., De Kauwe, M. G., Ukkola, A. M., Pitman, A. J., Guo, W., Hobeichi, S., & Briggs, P. R. (2021). Exploring how groundwater buffers the influence of heatwaves on vegetation function during multi-year droughts. *Earth Syst. Dynam*, *12*, 919–938. https://doi.org/10.5194/esd-12-919-2021

Ombadi, M., Nguyen, P., Sorooshian, S., & Hsu, K. lin. (2020). Evaluation of Methods for Causal Discovery in Hydrometeorological Systems. *Water Resources Research*, *56*(7). https://doi.org/10.1029/2020WR027251

Park, S., Kim, H., Jang, C., Zhang, L. T., & Martins, J. A. (2021). *Impact of Groundwater Abstraction on Hydrological Responses during Extreme Drought Periods in the Boryeong Dam Catchment, Korea*. https://doi.org/10.3390/w13152132

Paulik, C., Dorigo, W., Wagner, W., & Kidd, R. (2014). Validation of the ASCAT Soil Water Index using in situ data from the International Soil Moisture Network. *International Journal of Applied Earth Observation and Geoinformation*, *30*, 1–8. https://doi.org/10.1016/j.jag.2014.01.007

Petersen-Perlman, J. D., Aguilar-Barajas, I., & Megdal, S. B. (2022). Drought and groundwater management: Interconnections, challenges, and policyresponses. *Current Opinion in Environmental Science & Health*, *28*, 100364. https://doi.org/10.1016/J.COESH.2022.100364

Pezij, M., Augustijn, D. C. M., Hendriks, D. M. D., Weerts, A. H., Hummel, S., van der Velde, R., & Hulscher, S. J. M. H. (2019). State updating of root zone soil moisture estimates of an unsaturated zone metamodel for operational water resources management. *Journal of Hydrology X*, *4*. https://doi.org/10.1016/j.hydroa.2019.100040

Philip, S. Y., Kew, S. F., Van Der Wiel, K., Wanders, N., Jan Van Oldenborgh, G., & Philip, S. Y. (2020). Regional differentiation in climate change induced drought trends in the Netherlands. *Environmental Research Letters*, *15*(9), 094081. https://doi.org/10.1088/1748-9326/AB97CA

Qing, Y., Wang, S., Yang, Z. L., & Gentine, P. (2023). Soil moisture–atmosphere feedbacks have triggered the shifts from drought to pluvial conditions since 1980. *Communications Earth & Environment 2023 4:1*, *4*(1), 1–10. https://doi.org/10.1038/s43247-023-00922-2

Rigden, A. J., Mueller, N. D., Holbrook, N. M., Pillai, N., & Huybers, P. (2020). Combined influence of soil moisture and atmospheric evaporative demand is important for accurately predicting US maize yields. *Nature Food*, *7*, 20–28. https://doi.org/10.1038/s43016-020-0028-7

Rutger Weijers. (2020). *Drought indicators in The Netherlands*. http://repository.tudelft.nl/.

Schröder, J. J., Hilhorst, G. J., Oenema, J., Verloop, J., & van den Berg, W. (2021). *BOFEK2020 - Bodemfysische schematisatie van Nederland : update bodemfysische eenhedenkaart*. https://doi.org/10.18174/541544

Shi, H., Zhao, Y., Liu, S., Cai, H., & Zhou, Z. (2022). A New Perspective on Drought Propagation: Causality. *Geophysical Research Letters*, *49*(2). https://doi.org/10.1029/2021GL096758

Shiau, J. T. (2023). Causality-based drought propagation analyses among meteorological drought, hydrologic drought, and water shortage. *Science of The Total Environment*, *888*, 164216. https://doi.org/10.1016/J.SCITOTENV.2023.164216

Shorachi, M., Kumar, V., & Steele-Dunne, S. C. (2022). Sentinel-1 SAR Backscatter Response to Agricultural Drought in The Netherlands. *Remote Sensing*, *14*(10), 2435. https://doi.org/10.3390/RS14102435/S1

Siegmund, P. (2022, November 17). *KNMI - Four or six seasons?* https://www.knmi.nl/over-het-knmi/nieuws/vier-of-zes-seizoenen

Smith, L. A. (1988). INTRINSIC LIMITS ON DIMENSION CALCULATIONS. *PHYSICS LETIERS A*, *133*(6).

Su, Z., He, Y., Dong, X., & Wang, L. (2017). *Drought Monitoring and Assessment Using Remote Sensing* (pp. 151–172). https://doi.org/10.1007/978-3-319-43744-6_8

Sugihara, G., May, R., Ye, H., Hsieh, C. H., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, *338*(6106), 496–500. https://doi.org/10.1126/SCIENCE.1227079

Taylor, C. M., De Jeu, R. A. M., Guichard, F., Harris, P. P., & Dorigo, W. A. (2012). Afternoon rain more likely over drier soils. *NERC Open Research Archive*. https://doi.org/10.1038/nature11377

Tian, J., Han, Z., Bogena, H. R., Huisman, J. A., Montzka, C., Zhang, B., & He, C. (2020). Estimation of subsurface soil moisture from surface soil moisture in cold mountainous areas. *Hydrology and Earth System Sciences*, *24*(9), 4659–4674. https://doi.org/10.5194/hess-24-4659-2020

Tian, J., & Zhang, Y. (2023). Comprehensive validation of seven root zone soil moisture products at 1153 ground sites across China. *International Journal of Digital Earth*, *16*(2), 4008–4022. https://doi.org/10.1080/17538947.2023.2261902

Tu, L. (2019). *Downscaling SMAP Soil Moisture Data Using MODIS Data*. https://digitalcommons.lsu.edu/gradschool_theses

UN. (2022, November 15). *World Population Prospects 2022: Summary of Results | Population Division*. https://www.un.org/development/desa/pd/content/World-Population-Prospects-2022

Van Den Hurk, B., Siegmund, P., & Klein Tank, A. (2014). *KNMI'14: Climate Change scenarios for the 21st Century-A Netherlands perspective*. www.climatescenarios.nl

Van Der Velde, R., Benninga, H.-J. F., & Retsios, B. (2022). Twelve years profile soil moisture and temperature measurements in Twente, the Netherlands. *Open Access Earth System Science Data Discussions* . https://doi.org/10.17026/dans-xr2-m6d8

Van Der Velde, R., Colliander, A., Pezij, M., Benninga, H. J. F., Bindlish, R., Chan, S. K., Jackson, T. J., Hendriks, D. M. D., Augustijn, D. C. M., & Su, Z. (2021). Validation of SMAP L2 passive-only soil moisture products using upscaled in situ measurements collected in Twente, the Netherlands. *Hydrology and Earth System Sciences*, *25*(1), 473–495. https://doi.org/10.5194/HESS-25-473-2021

Van Hateren, T. C., Chini, M., Matgen, P., & Teuling, A. J. (2021). *Ambiguous Agricultural Drought: Characterising Soil Moisture and Vegetation Droughts in Europe from Earth Observation*. https://doi.org/10.3390/rs13101990

Van Wijnen, M., & Schotting, R. J. (2019). *Fresh water-brackish water interface monitoring in Vitens' groundwater extraction areas*.

Verdonschot, & Piet. (2017). *Integraal natuurherstel in beekdalen Ontwikkeling van diffuse afvoersystemen, gedempte afvoerdynamiek en beekprofielherstel*.

Wagner, W., Lemoine, G., & Rott, H. (1999). A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data. *Remote Sensing of Environment*, *70*(2), 191–207. https://doi.org/10.1016/S0034-4257(99)00036-X

Wakigari, S. A., & Leconte, R. (2022). *Enhancing Spatial Resolution of SMAP Soil Moisture Products through Spatial Downscaling over a Enhancing Spatial Resolution of SMAP Soil Moisture Products through Spatial Downscaling over a Large Watershed: A Case Study for the Susquehanna River Basin in the Northeastern United States*. https://doi.org/10.3390/rs14030776

Wanders, N., Henny Van Lanen, & Anne Van Loon. (2010). *Indicators for drought characterization on a global scale*. https://www.researchgate.net/publication/254843932

Wang, C., Fu, B., Zhang, L., & Xu, Z. (2019). Soil moisture–plant interactions: an ecohydrological review. *Journal of Soils and Sediments*, *19*(1), 1–9. https://doi.org/10.1007/S11368-018-2167-0/METRICS

Wang, L., & Chen, W. (2014). Equiratio cumulative distribution function matching as an improvement to the equidistant approach in bias correction of precipitation. *Atmospheric Science Letters*, *15*(1), 1–6. https://doi.org/10.1002/ASL2.454

Wang, L., Zhang, X., Wang, S., Salahou, M. K., & Fang, Y. (2020). Analysis and Application of Drought Characteristics Based on Theory of Runs and Copulas in Yunnan, Southwest China. *Environmental Research and Public Health*. https://doi.org/10.3390/ijerph17134654

Wang, Y., Yang, J., Chen, Y., De Maeyer, P., Li, Z., Duan, W., Wang, Y., Yang, J., Chen, Y., De Maeyer, P., Li, Z., & Duan, W. (2018). Detecting the Causal Effect of Soil Moisture on Precipitation Using Convergent Cross Mapping. *NatSR*, *8*(1), 12171. https://doi.org/10.1038/S41598-018-30669-2

Weng, Z., Niu, J., Guan, H., & Kang, S. (2023). Three-dimensional linkage between meteorological drought and vegetation drought across China. *Science of The Total Environment*, *859*, 160300. https://doi.org/10.1016/J.SCITOTENV.2022.160300

Wilhite, D. A., Glantz, M. H., & And Glantz, M. H. (1985). *Understanding the Drought Phenomenon: The Role of Definitions*. http://digitalcommons.unl.edu/droughtfacpubhttp://digitalcommons.unl.edu/droughtfacpub/20

WMO. (2012). *Standardized Precipitation Index User Guide*.

Wu, J., Miao, C., Zheng, H., Duan, Q., Lei, X., & Li, H. (2018). Meteorological and Hydrological Drought on the Loess Plateau, China: Evolutionary Characteristics, Impact, and Propagation. *AGU100*. https://doi.org/10.1029/2018JD029145

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

Xu, Z., Man, X., Duan, L., & Cai, T. (2022). Improved Subsurface Soil Moisture Prediction from Surface Soil Moisture through the Integration of the (de)Coupling Effect. *Journal of Hydrology*, *608*.

Xu, Z., Wu, Z., Shao, Q., He, H., & Guo, X. (2023). From meteorological to agricultural drought: Propagation time and probabilistic linkages. *Journal of Hydrology: Regional Studies*, *46*, 101329. https://doi.org/10.1016/J.EJRH.2023.101329

Yang, X., Lu, M., Wang, Y., Wang, Y., Liu, Z., Chen, S., Toscano, S., Franzoni, G., & Álvarez, S. (2021). Response Mechanism of Plants to Drought Stress. *Horticulturae* . https://doi.org/10.3390/horticulturae7030050

Zeng, Y., Hao, D., Huete, A., Dechant, B., Berry, J., Chen, J. M., Joiner, J., Frankenberg, C., Bond-Lamberty, B., Ryu, Y., Xiao, J., Asrar, G. R., & Chen, M. (2022). Optical vegetation indices for monitoring terrestrial ecosystems globally. *Nature Reviews Earth and Environment*, *3*(7), 477–493. https://doi.org/10.1038/s43017-022-00298-5

Zhang, H., Ding, J., Wang, Y., Zhou, D., & Zhu, Q. (2021). Investigation about the correlation and propagation among meteorological, agricultural and groundwater droughts over humid and arid/semi-arid basins in China. *Journal of Hydrology*, *603*, 127007. https://doi.org/10.1016/J.JHYDROL.2021.127007

Zhang, Q., Lv, X., Yu, X., Ni, Y., Ma, L., & Liu, Z. (2022). Species and spatial differences in vegetation rainfall interception capacity: A synthesis and meta-analysis in China. *CATENA*, *213*, 106223. https://doi.org/10.1016/J.CATENA.2022.106223

Zhang, W., Wang, Z., Lai, H., Men, R., Wang, F., Feng, K., Qi, Q., Zhang, Z., Quan, Q., & Huang, S. (2023). Dynamic Characteristics of Meteorological Drought and Its Impact on Vegetation in an Arid and Semi-Arid Region. *Water (Switzerland)*, *15*(22). https://doi.org/10.3390/w15223882

Zhang, X., Hao, Z., Singh, V. P., Zhang, Y., Feng, S., Xu, Y., & Hao, F. (2022). Drought propagation under global warming: Characteristics, approaches, processes, and controlling factors. *Science of The Total Environment*, *838*, 156021. https://doi.org/10.1016/J.SCITOTENV.2022.156021

Zhang, Z., Ma, J., Wang, T., Song, W., Hao, · Lu, & Hao, L. (2023). *Identify the relationship of meteorological drought and ecohydrological drought in Xilin Gol Grassland, China*. *116*, 2549–2564. https://doi.org/10.1007/s11069-022-05778-x

Zhao, J., Huang, S., Huang, Q., Wang, H., Leng, G., & Fang, W. (2020). Time-lagged response of vegetation dynamics to climatic and teleconnection factors. *CATENA*, *189*, 104474. https://doi.org/10.1016/J.CATENA.2020.104474

Zhuang, R., Manfreda, S., Zeng, Y., Romano, N., Dor, E. Ben, Maltese, A., Nasta, P., Francos, N., Capodici, F., Paruta, A., Ciraolo, G., Szabó, B., Mészáros, J., Petropoulos, G. P., Zhang, L., & Su, Z. (2021). UAS Based Soil Moisture Downscaling Using Random Forest Regression Model. *EGU21*. https://doi.org/10.5194/EGUSPHERE-EGU21-15569

Zhuang, R., Zeng, Y., Manfreda, S., & Su, Z. (2020). Quantifying long-term land surface and root zone soil moisture over Tibetan plateau. *Remote Sensing*, *12*(3). https://doi.org/10.3390/rs12030509

10 APPENDIX A

10.1 Data Management Plan (DMP)

10.1.1 Research Data Collection

Table S1: RDM considerations, best practices, and planned strategies.

1. NAME OF DATA FILE	2. SOURCE (PRIMARY OR SECONDARY DATA)	3. IF SECONDARY, WHO IS THE OWNER?	4. RESTRICTIONS AND LICENSE	5. DATA FORM	6. DATA FORMAT	7. CONTAINS PERSONAL DATA (Yes/No)
Global Surface Soil Moisture (1KmSSM)	Secondary data Earth Engine Data Catalog	(Han et al., 2023)	No license available	Imagery	.tiff	No
ECMWF Reanalysis v5 (ERA5)	Secondary data Earth Engine Data Catalog	European Centre for Medium-Range Weather Forecasts (ECMWF)	License available here.	Imagery	GRIB	No
Netherlands Soil Map	Secondary data Netherlands Data Portal	(Schröder et al., 2021)	License available here.	Vector	.shp	No
Actuell Hoogetebestand Nederland (AHN3)	Secondary data Earth Engine Data Catalog	Actueel Hoogtebestand Nederland	No license available	Lidar Imagery	.tiff	No
MOD09GA version 6.1	Secondary data Earth Engine Data Catalog	United States Geological Survey (USGS)	No license available	Satellite Imagery	.tiff	No
Twente Soil Moisture	Secondary data Source	(Van Der Velde et al., 2022)	Creative Commons, CC BY 4.0 license	Spreadsheet	.csv	No
Netherlands Meteorological data	Secondary data Source	The Koninklijk Nederlands Meteorologish Instituut (KNMI)	Creative Commons, CC BY 4.0 license	Spreadsheet	.csv	No
Groundwater measurements	Secondary data Source	Dinoloket	Creative Commons, CC BY 4.0 license	Spreadsheet	.csv	No

10.1.2 Organizing and Documenting Research Data

Table 6: RDM considerations, best practices, and planned strategies.

Name of main folder(s):	Msc_Thesis_Data_2024/06/24
Name of secondary and tertiary folders (if applicable):	Secondary: Objective_1, Objective_2, Objective_3 and Objective_4
Version control strategy:	Raw Data: "Objective_n_RawData_YYYYMMDD" Processed Data: " Objective_n_ProcessedData" Analysis Script: " Objective_n_AnalysisScript"
Metadata standards used:	
Readme file contents:	

10.1.3 Storage and Sharing of Research Data

Table 7: RDM considerations, best practices, and planned strategies.

Data	**Storage locations**	**Back-up location and frequency**	**Strategy to prevent unauthorized access to data during research**	**Strategy for pseudonymization or anonymization of data (if applicable)**
Master files	On my laptop's local drive. This ensured convenience during analysis.	Google Drive.	Access was regulated through strong passwords and fingerprints	
Copy of the data	My external Drive. For easy access.	My external Drive	Access was regulated through strong passwords and fingerprints	

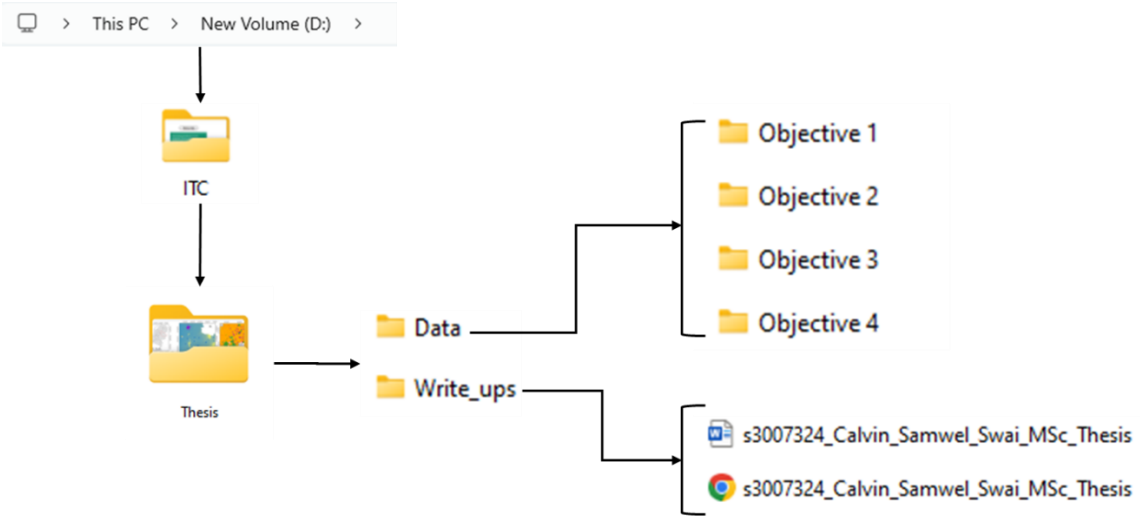


Figure S9: Data Storage and Organization

10.1.4 Research Data Long-Term Storage for Reuse

Table 8: RDM considerations, best practices, and planned strategies.

Names of data sets selected for long-term storage.	Groundwater_Datasets RZSM_estimates
Long-term storage option selected and time for data preservation:	Long-term
Preferable file formats used or to be used for long-term storage.	Preferred Format: Tiff, CSV (Comma-Separated Values) Example: Thesis_objective_1_sm_2011-2020.CSV Thesis_objective_3_NDWI_January_2020.CSV
Long-term data access (public/restricted) and why:	Public Access: I strongly believe in open science. By sharing my data with everyone, I hope to speed up the study and deepen our understanding of drought dynamics. This allows many different researchers to check, verify, and build on my work. This not only makes sure my work is accurate and can be repeated, but also helps science to move forward including the implementation of the recommendations derived from my research.

10.2 Other Figures

The capacity of ERA5 to reproduce KNMI-measured precipitation anomalies was assessed by comparing SPI-1 computed using both ERA5 and KNMI stations. Figure S1 illustrates that SPI computed with ERA5 effectively reproduced the time series of the SPI index computed from KNMI precipitation measurements. With RMSE below 0.5 and correlation score between 0.82 and 0.86, SPI computed using ERA5 for the period between 2000-2020 exhibited both similarity and close linear relation with to SPI computed using KNMI precipitation measurements.

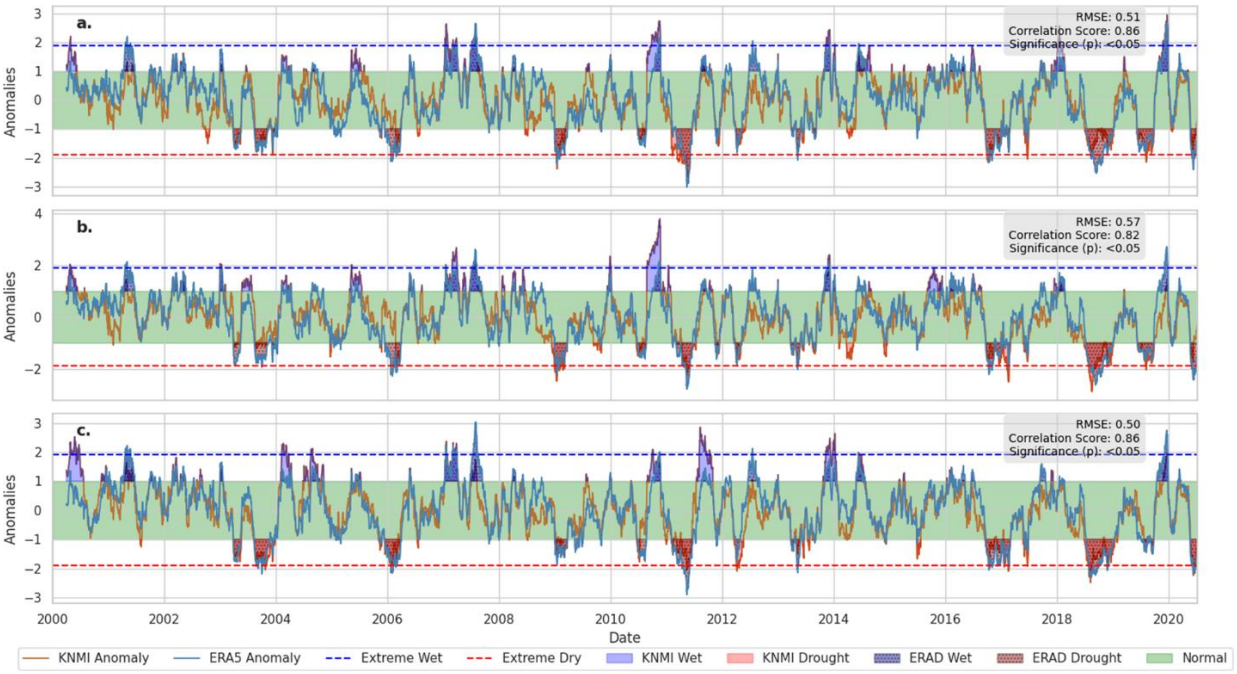


Figure S1: KNMI vs ERA5 Anomalies for KNMI stations a) Twente, b) Hupsel, and c) Heino meteorological stations.

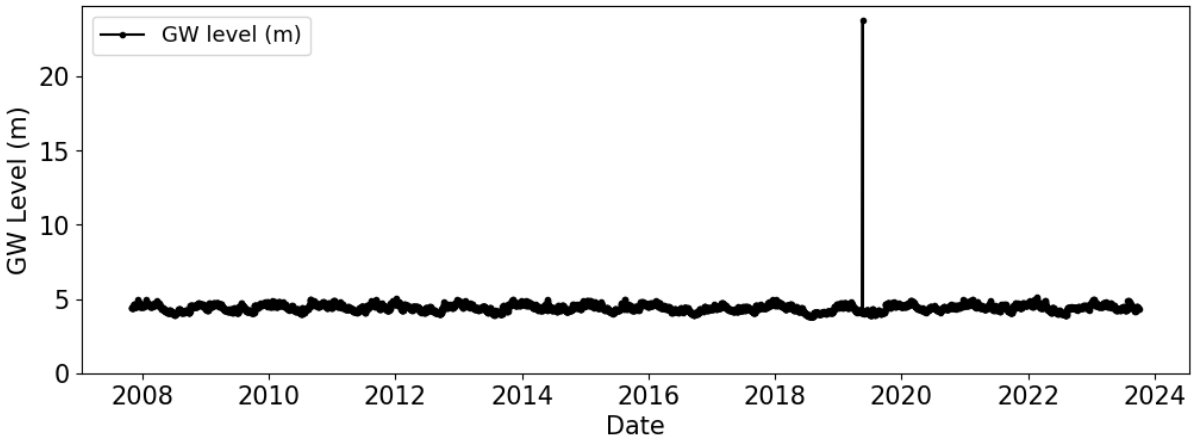


Figure S2: Groundwater level time series of the B22C0675 monitoring well showing outlier in the time series.

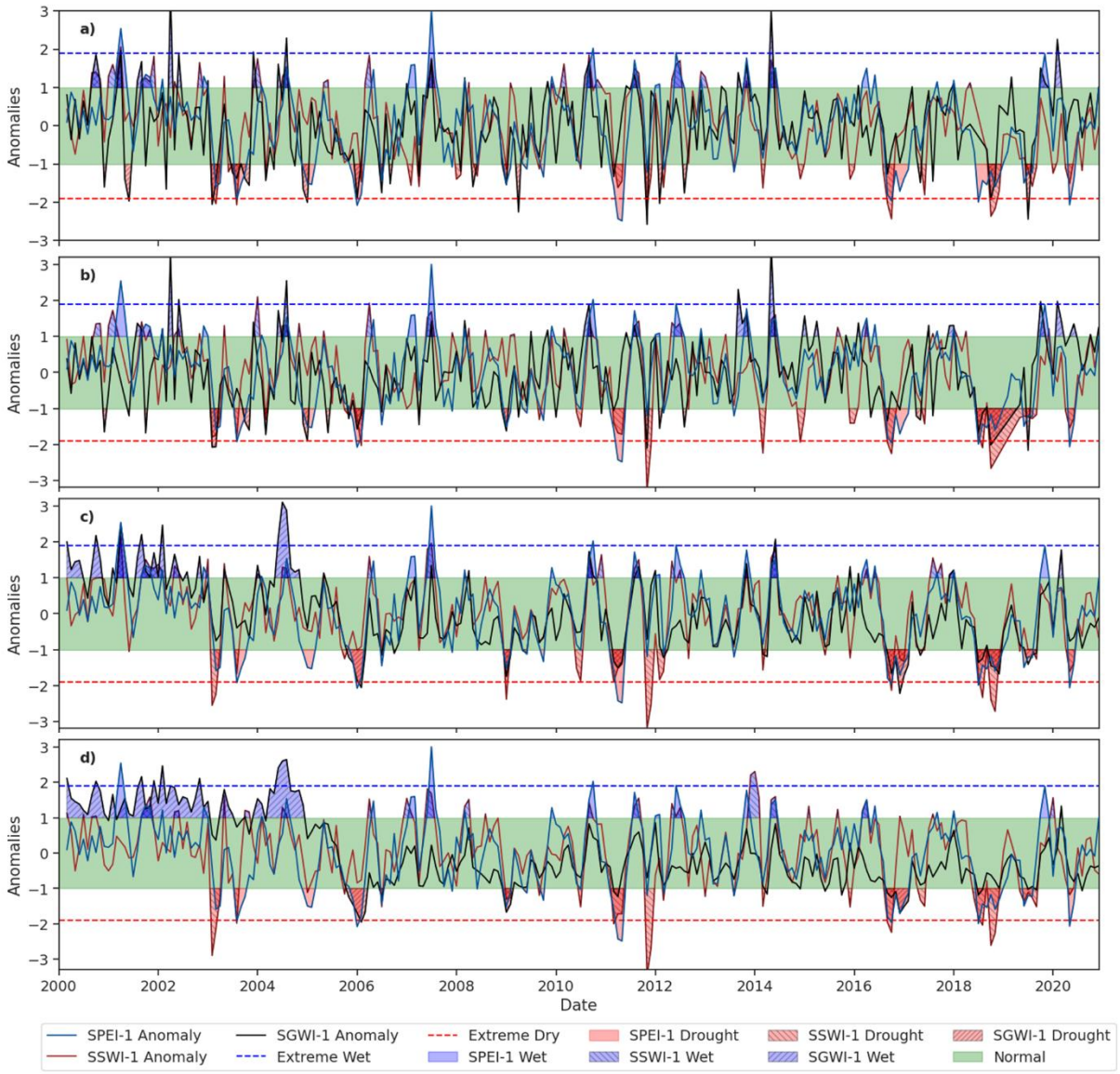


Figure S3: Temporal pattern of SPEI-1, SSWI-1, SGWI-1 for monitoring wells A) B22C0107, B) B22C0086, C) B22C0131 and D) B22C0089.

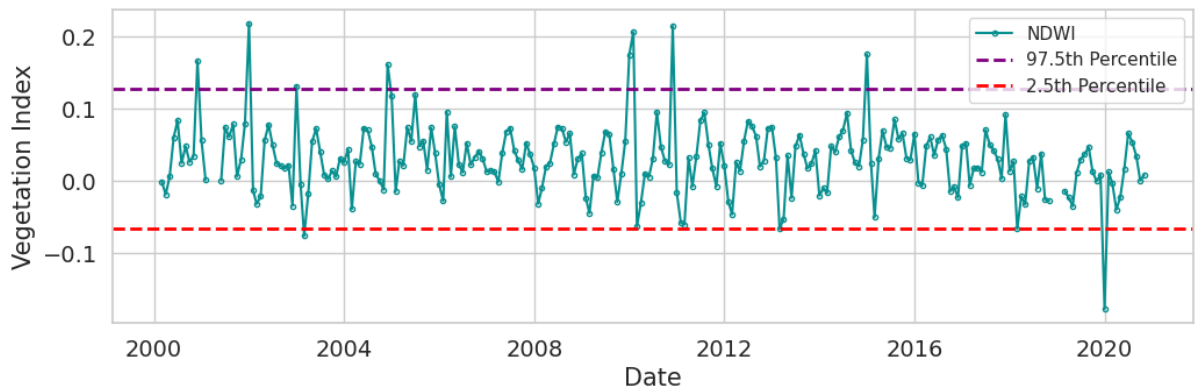


Figure S4: Temporal pattern of median NDWI of the Witharen showing extreme values in the time series.

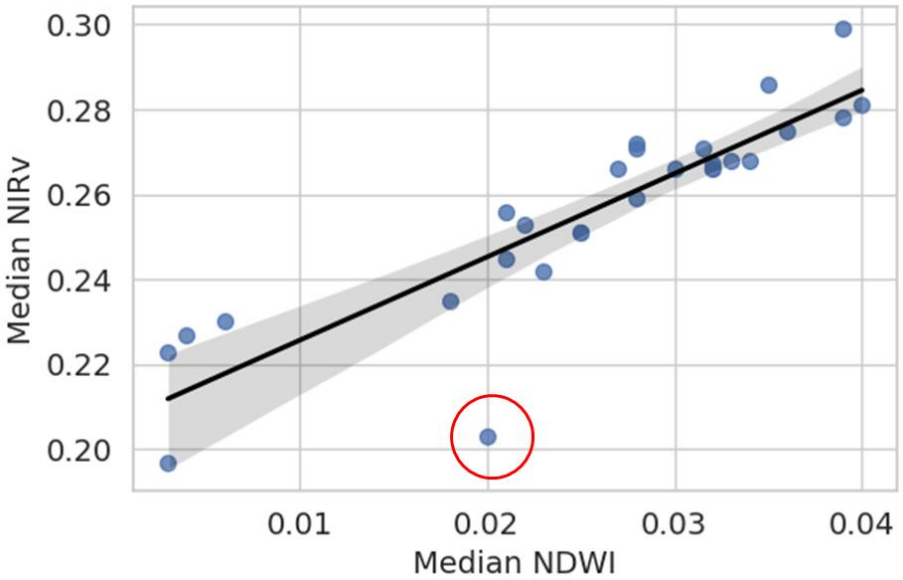


Figure S5: A scatterplot of the median NDWI and NIRv showing deviation of the monitoring well B22C0136's NDWI from the linear relationship.

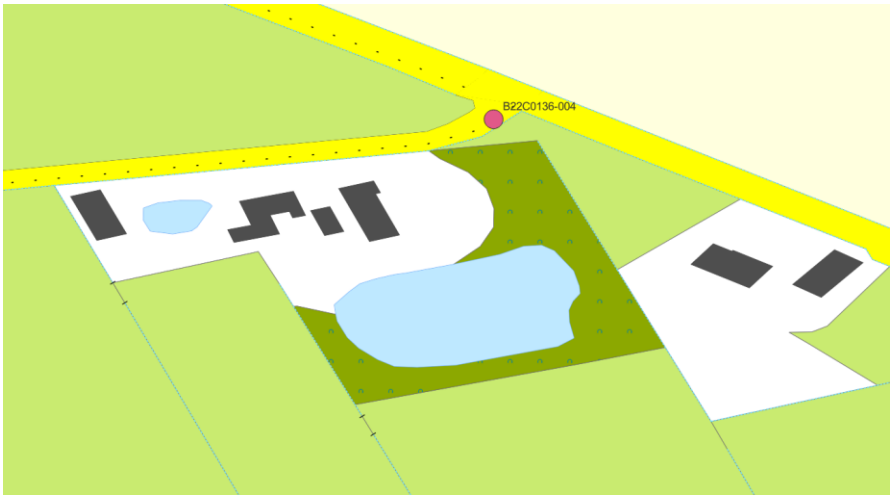


Figure S6: A snapshot of the Top10NL land cover properties around the B22C0136 monitoring well.

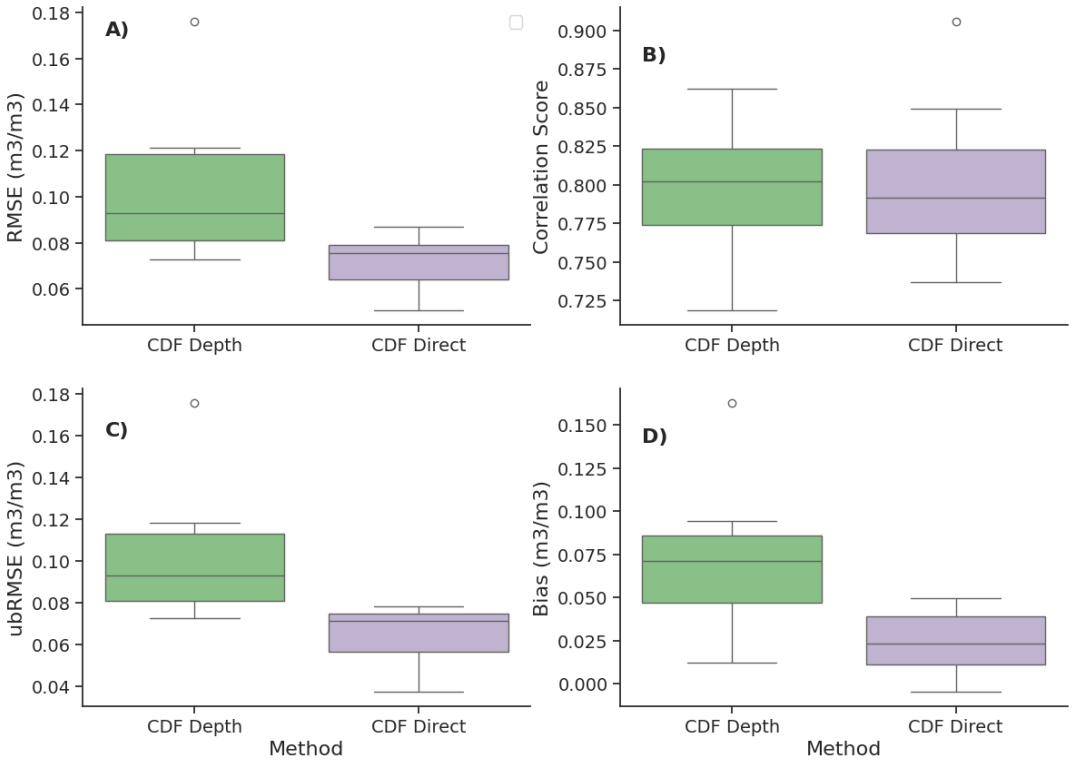


Figure S6: Performance of the two CDF techniques i.e., CDF applied directly to the $RZSM_{in\text{-}situ}$ (CDF Direct) and the one which was applied separately on each depth.

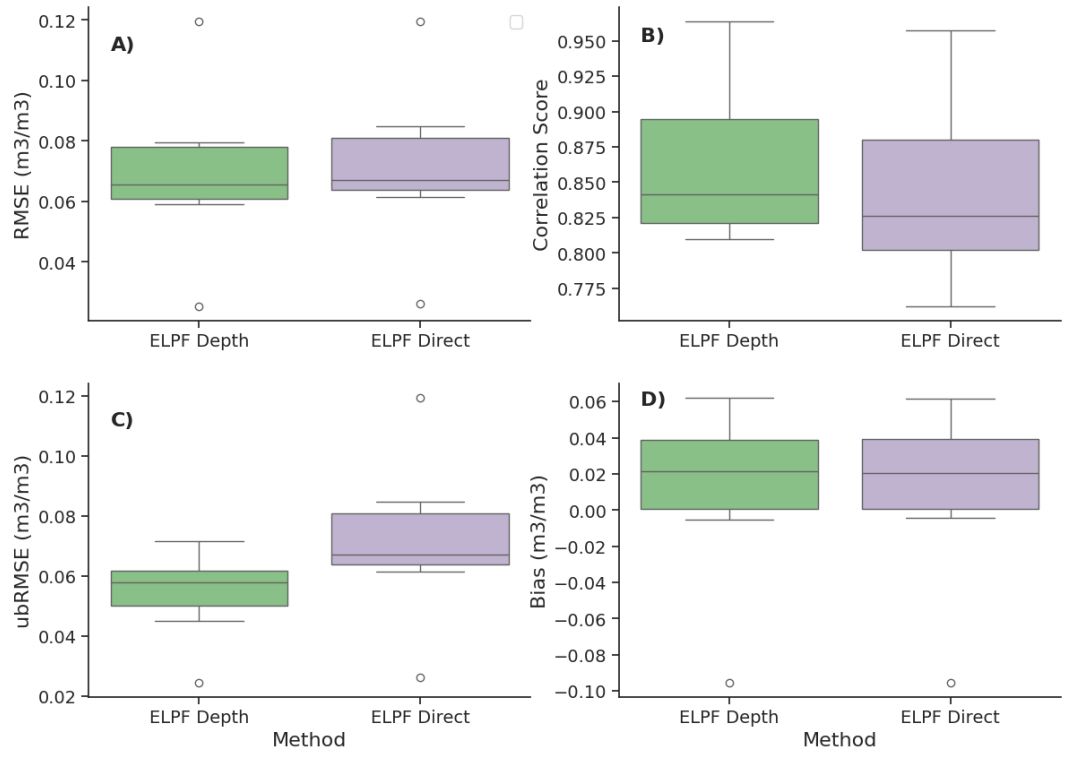


Figure S7: Performance of the two CDF techniques i.e., CDF applied directly to the $RZSM_{in\text{-}situ}$ (CDF Direct) and the one which was applied separately on each depth.

Figure S8: Snapshot of the detailed land cover terrain of the Witharen region. Source Top10NL (WMS:https://service.pdok.nl/brt/top10nl/wms/v1_0).