Agricultural and Forest Meteorology 341 (2023) 109683

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Dryspells and Minimum Air Temperatures Influence Rice Yields and their Forecast Uncertainties in Rainfed Systems

Abhijeet Abhishek [a], Mantha S. Phanikumar [a,b], Alicia Sendrowski [a], Konstantinos M. Andreadis [c], Mahya G.Z. Hashemi [a], Susantha Jayasinghe [d], P.V. Vara Prasad [e], Roberts J. Brent [f], Narendra N. Das [a,g,*]

[a] Department of Civil and Environmental Engineering, Michigan State University, East Lansing, MI 48824, USA
[b] MSU AgBioResearch, East Lansing, MI 48824, USA
[c] Department of Civil and Environmental Engineering, University of Massachusetts Amherst, Amherst, MA 01002, USA
[d] Asian Disaster Preparedness Center, Bangkok 10400, Thailand
[e] Department of Agronomy, Kansas State University, Manhattan, KS, 66506, USA
[f] NASA Marshall Space Flight Center, Huntsville, AL 35805, USA
[g] Department of Biosystems and Agricultural Engineering, Michigan State University, East Lansing, MI 48824, USA

ARTICLE INFO

Keywords:
Hydrologic-crop modeling
Climate forecast ensembles
Drought
Rice
Crop yield uncertainty
Cambodia
Mekong
Growing season variability
Mutual information

ABSTRACT

Crop yield predictions on inter-seasonal to inter-annual horizons are subject to a diverse set of uncertainties associated with climate forecast scenarios. However, the uncertainties associated with climate-related parameters can be marginally controlled (reduced) over the growing season, i.e., from planting through harvest, leveraging a suite of climate forecast ensembles as forcings. In this study, we present a novel approach that combines a coupled hydrologic-crop modeling framework with probabilistic forecasts to characterize and reduce uncertainties in seasonal rice yields predictions. By weighting real-time (nowcast) and forecast climate forcings, the comprehensive framework accurately quantifies uncertainties associated with climate forecasts. At a provincial scale, the crop model extensively captured the uncertainties in yields whilst significantly complementing observations over the growing season. We observed that the spread of yield predictions gradually decreased over time (i.e., towards harvest) as subsequent timeframes incorporated a higher degree of present conditions/forcings and reduced reliance on forecasts. Furthermore, we investigated the information exchange between yields and hydrologic/drought variables over different time frames within the season. Notably, we found a higher synchronization of information transfer between yields, dryspells, and minimum air temperatures towards the end of season, indicating strong explicit links between these variables and crop yields. These outcomes have significant implications on crop yield forecasting and nowcasting, particularly in data-poor regions. By providing a better understanding of the uncertainties associated with seasonal climate forecasts and the interplay between hydrologic variables, drought conditions, and crop yields, this research can aid in improving decision-making processes related to agricultural planning, and risk management. Moreover, these insights can inform assessments of economic, social, and environmental impacts of drought in agricultural systems.

1. Introduction

Emerging megatrends such as population growth, land-use changes, resource scarcity, and socio-economic practices profoundly impact agricultural ecosystems around the world. One of the most pressing challenges is the changing climate, including rising temperatures, untimely and sporadic precipitation patterns, shifting seasonality, and increased vulnerability to extreme weather events such as droughts and floods (Alizadeh et al., 2020; Clarke et al., 2022; Dai, 2013; Gampe et al., 2021). These changes in climate are multifaceted and pose a serious threat to agricultural sustainability and food security (Challinor, 2011; Fujimori et al., 2019; Jägermeyr et al., 2021). The situation is particularly acute in food insecure (developing) regions, for e.g., the Sub-Saharan Africa and Southeast Asia (Mekong Basin), wherein a combination of escalating natural (frequent exposures to extreme events such as hurricanes/cyclones, heatwaves, drought, etc.) and

* Corresponding author.
E-mail address: dasnaren@msu.edu (N.N. Das).

https://doi.org/10.1016/j.agrformet.2023.109683
Received 17 February 2023; Received in revised form 18 August 2023; Accepted 21 August 2023
Available online 31 August 2023

anthropogenic (e.g., physiographic and demographic shifts, conversion of forests and grasslands into croplands, etc.) activities is expected to further aggravate the current situation/crisis in these climate-sensitive hotspots, especially for food production (Ahmadalipour et al., 2019; Kang et al., 2021; Konapala et al., 2020; Padrón et al., 2020; Thilakarathne and Sridhar, 2017).

Climate and weather conditions significantly influence crop growth and development, underscoring the importance of accurate and timely weather forecasting for effective agricultural planning. Seasonal climate forecasts provide crucial insights for rainfed agricultural systems, aiding farmers in optimizing practices based on expected climate patterns (Lal and Stewart, 2018). Timely access to critical weather information, such as rainfall and temperature variability, enables informed decisions, optimizing practices from planting to post-harvest activities. Particularly, farmers can adjust strategies based on expected patterns selecting drought-tolerant varieties during anticipated dry periods or aligning harvest schedules with impending extreme events. Furthermore, a reliable forecast and an understanding of climate-crop links enhance productivity and risk minimization. For instance, forecasted yield estimates allow farmers to assess production and income risks, prompting adaptive strategies like crop insurance and diversification. Likewise, anticipating market conditions enables informed decisions on timing, volume, and pricing, fostering stability. Numerous earlier studies have demonstrated that improved yield forecasts enhance farmer's decision-making abilities (Australia: Brown et al., 2018; India: Kushwaha et al., 2022; United States: Lacasa et al., 2023). In contrast, unreliable forecasts can result in suboptimal practices such as inadequate water management, inaccurate fertilization, incorrect planting dates. Although, uncertainties in seasonal climate forecasts, stemming from meteorological influences and/or model intricacies, are inevitable, identifying the sources of uncertainties in seasonal climate forecasts, and implementing mitigation strategies are vital for effective decision-making and ensuring sustainable crop productivity.

Extensive research has been conducted to understand the relationship between climate and crop yields, with particular focus on the impacts of climate variability on staple crops (e.g., maize, rice, wheat, soybeans, etc.) (Innes et al., 2015; Mavromatis, 2015). As outlined in many of these studies, there is compelling evidence of the adverse effects of changing climatic conditions on interannual crop yields, thus emphasizing the urgent need to address this burgeoning crisis of food security (Minoli et al., 2022; Vogel et al., 2019; Rifai et al., 2019; Shukla et al., 2019). Shortened growing seasons, altered precipitation patterns, warmer growing season temperatures (i.e., the Growing Degree Days, GDD), and increased intensity and frequency of flood and drought occurrences threaten water availability and agricultural productivity in developing agrarian nations (Challinor et al., 2016; Ray et al., 2015; Zhao et al., 2017). For instance, Ortiz-Bobea et al. (2021) observed a temperature-related linear decline in agricultural productivity over the growing season, varying across countries. Hence, the existing literature underscores the urgency of addressing the impacts of climate variability on agricultural crop yields through sustainable and climate-resilient approaches.

Skillful forecasts of crop yields at seasonal and sub-seasonal timescales are essential for facilitating improved crop management and decision-making given the myriad uncertainties in agricultural production (Basso and Liu, 2019; Togliatti et al., 2017). Weekly to monthly yield projections not only ensure food security but also provide an overview of the weather conditions in the adjacent growing season. Such critical pieces of information can allow governments, research institutions, and policymakers, to act accordingly, and implement effective adaptation measures to build resilient agricultural systems (e.g., early warning systems for extreme events, water management techniques, etc.). Numerous previous studies have forecasted crop yields at varying spatial and temporal scales using process-based models and satellite remote sensing techniques (Karthikeyan et al., 2020; Konduri et al., 2020). However, relying on a combination of approaches—use of computational modeling, data analytics, automation, advisory systems, remote sensing, and appropriate agronomic practices—can yield better results compared to a single approach. Incorporating satellite remote sensing data into process-based models allows for a comprehensive assessment of crop conditions and facilitates targeted interventions. Generally, satellite imagery captures the vegetation dynamics, and crop attributes including leaf area index (LAI), land surface temperature at varied temporal and spatial coverages monitor the real-time crop growth and yield potential. Additionally, the weather forcings (e.g., precipitation, air temperature, etc.) that are provided to the models, assimilation of remotely sensed state variables (e.g., soil moisture), calibration and validation of the models, among others are some of the key components used to capture crop characteristics (e.g., phenology, yield, etc.). Similarly, combining satellite datasets with process-based models leads to improved simulation accuracy while enhancing mechanistic understanding of the physiological and biochemical processes of crop growth. Specifically, they track the crop development (phenology) over the growing season based on climate, soil properties, agronomic practices, etc. Likewise, the use of data (statistical) analytics (e.g., historic trends/patterns of cropping), and knowledge about local agronomic practices (e.g., planting/sowing date, fertilizer application, cultivars, etc.) enables accurate predictions, optimized resource allocation, whilst identifying yield gaps and inefficiencies.

An integrated approach leverages the strength of multiple techniques, thus allowing enhanced decision-making and improved efficiency. In other words, the integration of advanced modeling techniques with real-time climate data offers farmers the potential to enhance productivity, reduce risks, and improve resource management. Particularly, by combining hydrologic and crop models, we overcome the limitations of representing the complex hydrological processes solely within the crop models, and get a more accurate representation of water availability and drought conditions, which are essential factors influencing crop growth and productivity. The hydrologic model provides valuable insights on the soil moisture dynamics and water availability (compared to the simplified soil water balance module employed in crop models), enabling us to assess the water requirement for crop growth and development during periods of stress. Similarly, a combination of data analytics and automation enables the optimization of agricultural operations by analyzing large sets of data and leveraging technology to streamline processes, maximize productivity, and reduce wastage. Recently, several studies have incorporated new technologies such as machine learning, and deep learning specifically using neural networks, that sample the input variable space for a particular vegetation type (Ermida et al., 2017). These new systems boast sophisticated algorithms that provide the flexibility to analyze a large and diverse set of datasets, enabling a holistic understanding of agricultural systems. Such studies have exhibited improved prediction of interannual crop yields, when remote sensing information, and accurate agronomic management information, were supplemented with deep leaning and machine learning-based models (Feng et al., 2020; Paudel et al., 2023). Although integrated approaches demonstrate improved accuracy, their implementation is often hindered due to technical complexities, data availability and quality (e.g., inconsistent data), expensive computational (and maintenance) requirements, and contextual factors (e.g., differences in local agronomic practices).

While existing forecasting methods have advanced our understanding of crop yield dynamics, there are several limitations that hinder the accurate and timely estimation of yield projections, particularly in regions vulnerable to climate extremes. The Lower Mekong Basin (LMB), specifically Cambodia, serves as a compelling case study, as it has experienced significant impacts from weather extremities and anthropogenic factors, leading to adverse effects on rice production and the livelihoods of farmers (Guo et al., 2017; Hoang et al., 2019; Son et al., 2012; Thilakarathne and Sridhar, 2017). Hence, addressing these limitations requires advancements in data availability, modeling approaches, and stakeholder engagement.

Implementing the abovementioned approaches would greatly benefit in mitigating the impacts of such natural climatic shocks, whilst significantly reducing crop losses. Generally, the impact of climate-related variables that are subjected to a higher degree of uncertainty can be marginally modulated/controlled during the growing season through a multi-approach framework (Hansen et al., 2006). Hence, changes in crop phenology and development, especially from grain-filling until mature/harvest, are highly dependent on weather, so having a reliable long-term seasonal forecast of weather variables (e.g., precipitation and temperature) is key to achieving optimal yield forecasts. Generally, the meteorological data (input) forced into the model often contribute a significant proportion of uncertainty (Müller et al., 2021). This can be easily interpreted by referring to Fig. 1, which shows a considerable portion of uncertainty (highlighted in blue color) at the start of the growing season that gradually decreases over a period of time until it reaches crop maturity/harvest. Although uncertainties are inevitable, we used historic weather observations (as input) to characterize the uncertainty related to climate variability (represented by the white dashed line around the blue color in Fig. 1a). In this study, we focused on quantifying the uncertainties associated with climate inputs, whereas the uncertainties resulting from model parameters are expected to be covered in a forthcoming study that uses data assimilation (DA) to assess the variability in model uncertainties.

Additionally, we know that, forecast information can minimize the negative impacts of weather (i.e., drought stress) on crop phenology, thus resulting in lower crop losses and higher yields. Hence, recognizing the importance of such geophysical changes on specific sectors (e.g., agriculture, water resources) becomes imperative for understanding the complex web of relationships between these environmental processes and crop yields in the forecast mode. However, the dynamics of these large-scale geophysical systems are difficult to assess, often resulting in the underutilization of key information. In other words, the representation of physical processes and interactions among key variables in computational models are still not discerned and exploited to the full potential. This missing information can therefore have a large impact on understanding and quantifying the links between different geophysical parameters (Goodwell et al., 2020). Information Theory (IT) provides a comprehensive method to characterize the cause-and-effect relationships of large-scale geophysical processes through the extraction of information from single/multiple signal(s) (i.e., time series of a state variable) (Shannon, 1948). To date, a substantial amount of work has been done in hydrology using IT (Sendrowski and Passalacqua, 2017; Thiesen et al., 2019), however, no studies have implemented such an approach on an agro-hydrologic coupled system. As coupled systems have non-linear relationships between them, IT provides a systematic quantification of the missing links between various physical processes.

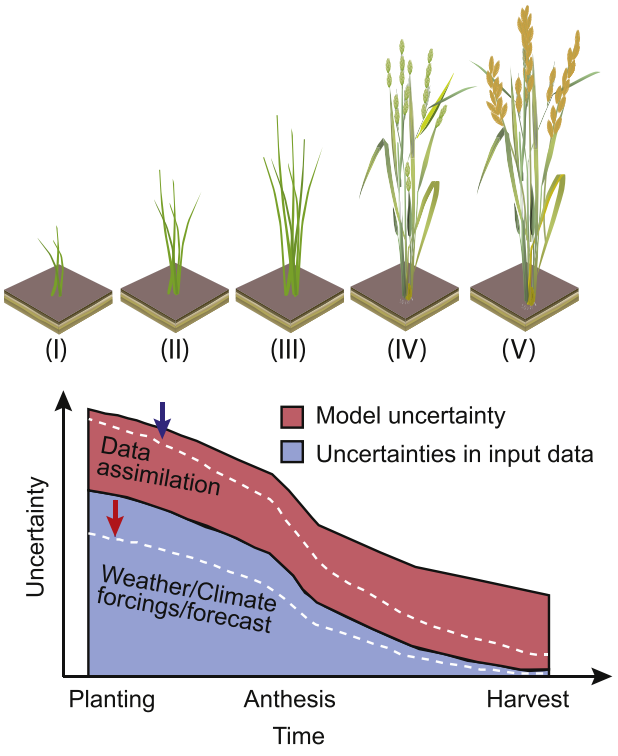

Fig. 1. Anatomy of crop yield prediction: The stacked plot (bottom left panel) exhibits the contribution of uncertainties, either from model parameters or input forcings, based on the various stages of crop (rice) growth: (I) Seedling, (II) Vegetative stages (Tillering, Stem elongation, Panicle Initiation), (III) Reproductive stage (Booting, Heading, Flowering), (IV) Grain filling, and (V) Maturity (top left panel). The white dashed line represents the reduction in model uncertainty (red) through data assimilation while the white dashed line in input climate forcings (blue color) represents the significant reduction in uncertainty using multi-model ensembles of seasonal climate forecast data, over the growing season, i.e., from planting through harvest.

The primary goal of our study was to evaluate the probabilistic seasonal rice yield forecasts (for the growing season) over Cambodia using the historic North American Multi-Model Ensemble (NMME) forecast datasets (Jha et al., 2019; Slater et al., 2019). We formulated the following research questions: 1) How do uncertainties in seasonal climate forecasts of crop yields evolve as the season progresses? and 2) What are the dominant hydrologic variables or derived indicators that influence crop yield forecasts through the growing season? Based on our questions, we hypothesized that seasonal climate forecast (independent variable) improves crop yield (dependent variable) predictions over time as more weather inputs (i.e., observations) are available (from planting through harvest). Thus, we implemented this using an integrated modular framework—the Regional Hydrologic Extremes Assessment System (RHEAS: Andreadis et al., 2017)—that provides an end-to-end composite hydrologic (and drought) and crop yield probabilistic nowcast (current) and seasonal forecast over Cambodia. Although RHEAS provides holistic real-time monitoring of drought and yield nowcasts covering a wide range of areas, this study is a significant extension to the current capabilities, wherein we examine the efficacy/skill of seasonal forecasts based on a multi-model ensemble forecast system. Specifically, we implemented a novel approach to estimating yields based on past weather observations (NMME) up to a certain forecast date, and weather data from other years are sampled for the rest of the season. In other words, the model integrates the observed weather data (from NMME) up to the current date with the rest of the input data sampled from climatology for the remaining season. In the end, we get an approximation of the climatic component of uncertainty (error) during each time window that represents the discrepancy between simulated and observed yields. In short, by using up-to-date climate information, we account for the temporal variability of climate variables during the growing season, capturing the effects of changing weather patterns on yield outcomes. Furthermore, we investigated the performance (impact) of associated drought/hydrologic (forecast) variables with crop yields by measuring the shared information between crop yield and hydrologic (and drought) variables. The combination of such statistical technique with hydrologic-crop modeling, and seasonal climate forecast sets this research apart, offering a more comprehensive and dynamic understanding of the factors influencing crop yield predictions. While it is widely known that uncertainties decrease over time, the novelty behind the approach lies in its comprehensive assessment and identification of key variables that primarily influence the uncertainties in seasonal climate forecasts of crop yields over the growing season. In other words, the study provides a detailed analysis of how uncertainties evolve (over time), and identify the key geophysical variables (i.e., hydrologic variables and derived drought indicators) that influence crop yield forecasts throughout the growing season. Following a comprehensive Introduction in Section 1, Section 2 describes the study area, datasets used, model architecture, and the methodology employed.

Section 3 presents the key results, followed by a discussion. Section 4 concludes the study by summarizing the key takeaways.

2. Materials and methods

2.1. Study area

Our study region, Cambodia (Fig. 2), has a diverse and unique ecological landscape, often characterized by a low-lying central plain that is encompassed by uplands and low mountains. The central plains host the vast freshwater lake Tonle Sap and the upper stretches of the Mekong River. The presence of the Tonle Sap and the Mekong River greatly influences the region's hydrology, agricultural practices, etc. The Mekong River and its tributaries (and floodplains) are crucial for irrigation, transportation, and fishing activities, serving as the lifeline for the local communities. During the wet season, the Mekong River swells, causing its tributaries and the Tonle Sap to reverse their flow. This inundation enriches the surrounding floodplains with nutrient-rich sediment, creating fertile conditions for agriculture. Likewise, the Tonle Sap plays a crucial role in supporting diverse ecosystems. The lake acts as a natural reservoir, expanding and contracting with the annual monsoon rains and the flow of the Mekong River (MRC, 2014). In addition, the central plains serve as the agriculturally fertile areas, with regions around Tonle Sap and the deltaic regions of the Mekong River hosting large areas of rice, sugarcane, cassava, and corn production. Rice is the primary grown crop in the region, with a large proportion of the population dependent on it for their livelihood (Mainuddin et al., 2013). Typically, rice is grown twice a year—a) wet season (influenced by the southwest monsoon), which lasts from May to October, and b) dry season (influenced by the northeast monsoon), which lasts from November to April. The annual precipitation ranges between 1200 and 2500 mm/year, with ~85% of the rainfall during the wet season (Lauri et al., 2014). Likewise, the average annual temperatures and evaporation rates range between 21 to 35°C and 1,000–2,000 mm/year respectively depending on the location (Liu et al., 2007). Overall, these prominent features greatly influence agriculture, hydrology, and the natural ecosystem in the region.

2.2. Composite modeling framework

In our study, we implemented an automated modular framework—the RHEAS—that integrates information from multiple components of the terrestrial water cycle (i.e., the soil-plant-atmosphere continuum) to effectively monitor the crop yields and corresponding hydrologic stresses. The system modularity of RHEAS framework enables it to handle an assortment of diverse datasets (see Table. S1) smoothly across multiple modeling systems, while facilitating the functioning of the models in tandem through a centrally located PostGIS (or PostgreSQL) spatially enabled relational database. Such design allows the seamless coupling of the models within the framework, thus making its design distinct and unique from other information systems. Specifically, RHEAS configuration follows the loose coupling of the Variable Infiltration Capacity (VIC: Liang et al., 1994) hydrologic model with a modified version of the Decision Support System for Agrotechnology Transfer (*m*-DSSAT: Jones et al., 2003) crop model. Fig. 3 illustrates the architecture of RHEAS framework in details. In practice, RHEAS has been successfully implemented in NASA-SERVIR hubs, such as East Africa and LMB, for operational drought and crop yield monitoring and forecasting (Abhishek et al., 2021; Andreadis et al., 2017). More information on RHEAS can be found at https://gitlab.com/kandread/RHEAS.

At the core of RHEAS, the VIC hydrological model captures the complex hydrological processes that govern the movement of water by incorporating spatially distributed information on various hydrological variables such as precipitation (PREC), minimum and maximum surface air temperature (TMIN and TMAX), evapotranspiration (ET), runoff, etc. Specifically, it simulates the hydrological variables and drought states (indices) over a gridded domain at a spatial resolution of 0.25° at a daily time-step in nowcast (i.e., estimation of current conditions) and forecast (i.e., estimation of future conditions) modes of operation. It incorporates the spatially distributed information on various hydrological variables to construct the key drought indices (or indicators), including the Standardized Precipitation Index (SPI), drought Severity (SEV), the Soil Moisture Deficit Index (SMDI), Dryspells, etc. The VIC hydrological model exhibits several advantages when integrated with other environmental models (e.g., a crop model in our case). Firstly, it solves the energy and water balance equations to provide a realistic estimation of water availability for crop growth and development. Secondly, it captures the dynamics of water movement through the plant-soil-atmosphere continuum, which is critical for understanding the impact of different hydrological processes on interannual crop yields. Furthermore, the VIC has been extensively tested and validated over different regions, ensuring its stability and reliability.

Although the VIC hydrologic model has traditionally been used for standalone real-time applications (Hamman et al., 2018), the modular framework of RHEAS makes it coherent for the VIC to interact with other crop growth models (e.g., DSSAT) via the PostgreSQL relational database system (Holzworth et al., 2015). Firstly, such a design allows the seamless flow of information and smooth handling of internal files/datasets between the constituent models (i.e., VIC and *m*-DSSAT),

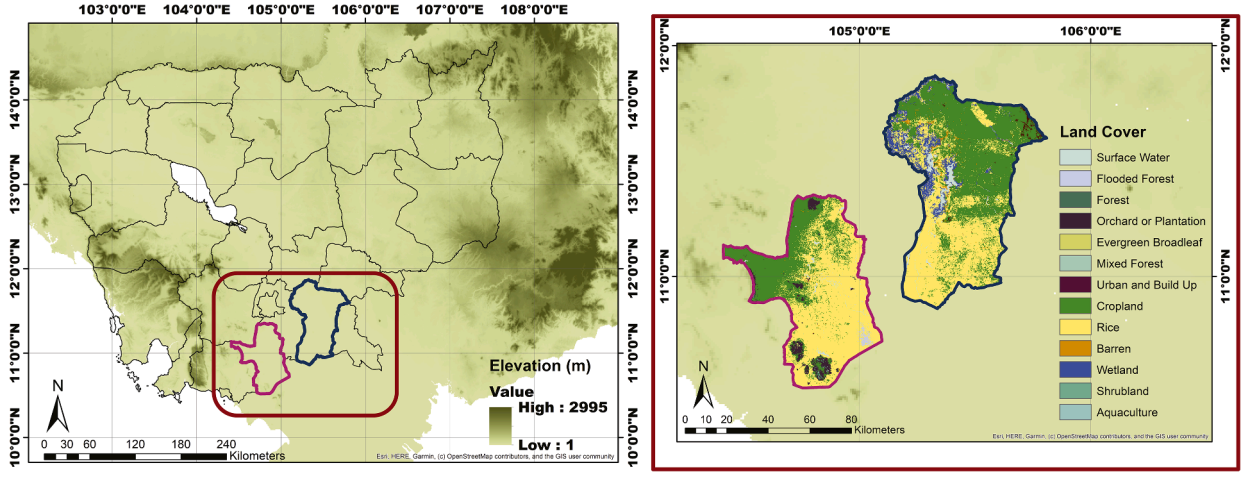

Fig. 2. Map showing the topography of Cambodia, with the two study areas—Takeo and Prey Veng provinces—highlighted in red and blue respectively. The inset shows the general landcover pattern over both provinces.

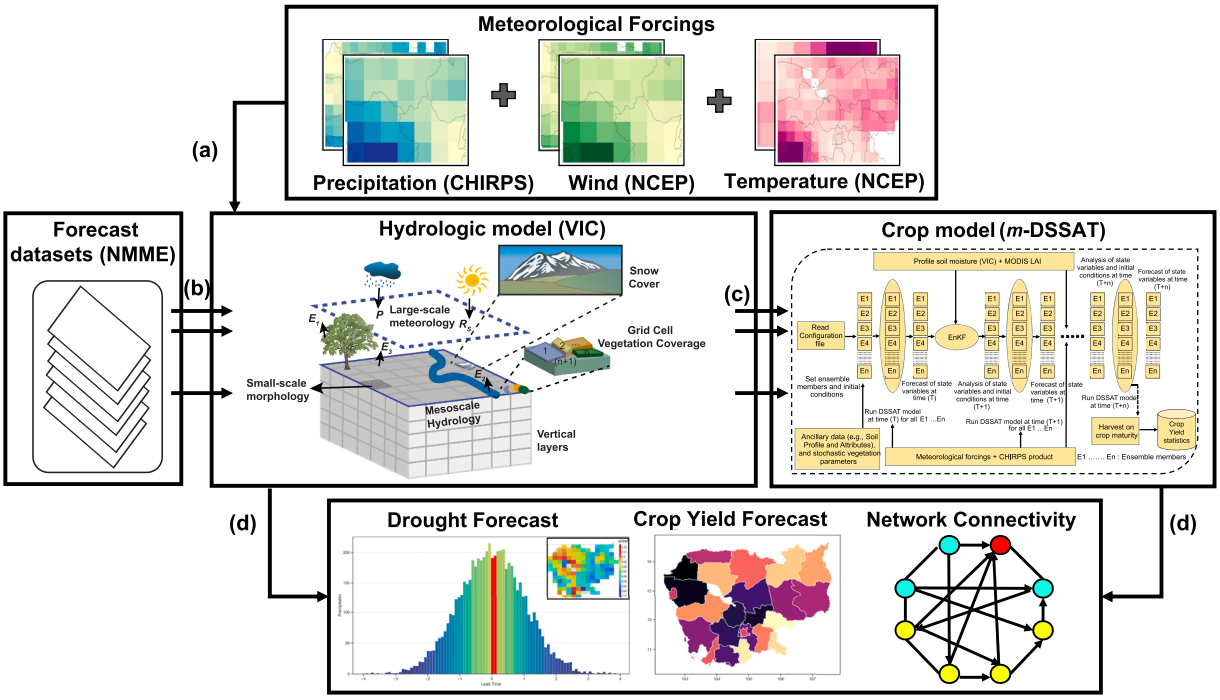

Fig. 3. Simplified illustration of the RHEAS framework. RHEAS follows a modular hybrid approach of loosely coupling a hydrologic model with a crop model. a, Using a set of meteorological forcings, the VIC hydrologic model simulates the core hydrologic variables and drought states in a nowcast (nowcast refers to the present conditions) mode; b, An ensemble of NMME datasets is next ingested into the model for seasonal forecast operations; c, Using the VIC-simulated outputs, the m-DSSAT crop model simulates an ensemble of crop yields, both in nowcast and forecast mode; d, Outputs of the model, specifically hydrologic variables (e.g., soil moisture) and drought indices (e.g., severity) are then used to construct the network connectivity/links between different components of the model outputs.

whilst yielding a composite hydrologic-agriculture data product. Secondly, the integration within RHEAS adds another layer of robustness to our approach, enabling us to make accurate predictions and assessments of crop yields under different climate and management scenarios.

Using the VIC-generated hydrologic variables (e.g., rainfall, solar radiation, surface air temperature, etc.), and other agronomic management information (cultivar genotype coefficients, fertilizers) as input, the *m*-DSSAT crop model simulates the growth, crop stage development, and end-of-season yield over the growing season. Due to the extensive testing and calibration using empirical data from various regions and cropping systems, DSSAT offers several advantages over other crop models (Kadiyala et al., 2015; Subash and Ram Mohan, 2012). Its inherent ability to incorporate information on agronomic practices, such as fertilizer application rates, cultivar genotype coefficients, irrigation rates, etc., enables the assessment of crop behavior under different management scenarios. Additionally, DSSAT has been extensively validated over different regions, including our study area (Abhishek et al., 2021) . However, unlike the traditional DSSAT, the *m*-DSSAT within RHEAS is a customized version that can run multiple ensembles member (~40 ensembles) for different soil properties and agronomic management practices (Ines et al., 2013).

Although DSSAT is widely used and recognized for its capability to simulate crop growth and yield, there are few notable limitations, such as the simplified soil water balance model employed within DSSAT, that may not accurately represent the complex hydrological processes. As a result, it can lead to potential discrepancies in simulating water availability and its impact on crop growth and yield. Therefore, to address this limitation, our study incorporates a hydrologic model within the RHEAS framework. By integrating the hydrologic model with the *m*-DSSAT crop model, we enhance the accuracy and robustness of our predictions. The hydrologic model captures the complex hydrological processes, including infiltration, evapotranspiration, and runoff, which are crucial factors influencing soil water availability and crop growth. Additionally, the VIC hydrological model simulates the land-atmosphere fluxes to capture the spatial and temporal dynamics of soil water availability over the growing season. Consequently, this improves the accuracy of crop yield predictions, as the hydrologic model accounts for the intricate interactions between water availability, plant physiological responses, and agronomic management practices.

2.3. Data

The high-resolution Climate Hazards Group Infrared Precipitation with Station data (CHIRPS: Funk et al., 2015) precipitation product and the National Centre for Environmental Prediction (NCEP: Kalnay et al., 1996) air temperature and wind speed gridded reanalysis product were used as meteorological/initial forcing into the hydrologic model. The CHIRPS product is a merged product that incorporates satellite-derived rainfall estimates with ground-based station data to produce a gridded precipitation dataset. Due to the availability of consistent (1981-present), high-resolution (0.05 degrees or ~5 km) blended products with extensive validation, it has been widely used for various applications, including agricultural monitoring, climate analysis, drought monitoring, and water resource management, especially in data-scarce regions (Katsanos et al., 2016; Toté et al., 2015). Likewise, the NCEP reanalysis product provides a comprehensive dataset of global atmospheric and oceanic observations. It utilizes a data assimilation approach to merge historical observations with model simulations, resulting in a consistent representation of the Earth's climate system over long periods. Thus, they have been widely used for climate research and weather analysis, contributing to our understanding of past climate conditions and aiding in the development of climate-related studies and applications (Eini et al., 2019; Hagemann and Dümenil Gates, 2001; Maurer et al., 2010). Additionally, the land cover information in the region was derived from the Moderate Resolution Imaging Spectroradiometer (MODIS-500 m) global product (Friedl et al., 2010), whilst the gridded soil information (10 km) was based on the SoilGrids1km by Han et al. (2019). A list of all the datasets used within the RHEAS

framework are summarized in Table. S1.

For the seasonal climate forecasts, we used the NMME seasonal prediction system that combines multiple climate models in an ensemble-based approach to provide skillful predictions of climate variables on seasonal to interannual timescales. The NMME dataset offers a comprehensive set of predictions, enabling the assessment of uncertainty and the exploration of future climate scenarios at a global scale. Examples of commonly predicted climate variables include surface air temperature, the net solar flux at the surface, specific humidity, precipitation, etc. All these variables are extensively evaluated and validated with observations to assess their skill and reliability. Thus, they have illustrated the potential in predicting the key climate phenomena such as El Niño-Southern Oscillation (ENSO), climate risk assessment on water resources and agriculture, etc. (Thober et al., 2015; Wang et al., 2022). Although the NMME data incorporates several state-of-the-art climate models, we used the climate forecasts from the Climate Forecast System version 2 (CFSv2) and the Community Climate System Model 4.0 (CCSM4) coupled modeling systems. Each model consists of multiple ensembles of climate prediction, with each ensemble based on different initial conditions, model physics, and data assimilation techniques. In practice, the multi-model ensemble approach has demonstrated exceptional skill/efficacy in quantifying prediction uncertainty and consistently outperforming individual model ensembles, thus making it an ideal choice for this study (Becker and van den Dool, 2016; Tippett et al., 2019). Specifically, we used the daily records of climate forecast data, particularly precipitation, and temperature, from both models in this study (discussed further in methods).

For validating the simulated crop yields, the actual yields (i.e., observations) were collected by the Department of Agricultural Land Resources Management, General Directorate of Agriculture, Phnom Penh, Cambodia. They provide food and agriculture information (including crop yield, fertilizer rates, etc.) over selected provinces in Cambodia and nearby areas by establishing mechanisms to collect data through field visits, surveys, or reports submitted by farmers. These data are aggregated at the provincial or national level, often involving quality control measures, including data validation, cross-checking with other sources, and addressing any inconsistencies or errors in the collected data. Other relevant information, such as the local agronomic practices (e.g., cultivar varieties, fertilizer rates, etc.), were obtained from published reports, local stakeholders, or agricultural databases (e.g., the Food and Agriculture Organization, FAO).

2.4. Methodology

Although RHEAS has been successfully implemented over large areas (such as the LMB) for assessing the spatial-temporal behavior of end-of-season rice yields and associated drought conditions, in this study, we configured our modeling domain over Cambodia (at a spatial resolution of 25 km). As is well-known, Cambodia has been at the forefront of frequent extreme events, such as droughts, that have adversely affected the agriculture sector, and the rural economy. Hence, in this study, our intent was to be region/country-specific, and we scaled down our analysis to provincial levels (often called administrative units or provinces). There are multiple reasons for doing so: 1) a number of previous studies have been carried out at the country level while there are hardly any district-specific studies, 2) country-level crop yields are very subjective and do not reflect the actual production/yield for major cropping districts, 3) evaluation of hydrologic/drought conditions at a district-level can reveal the outcome/performance of the final yields. Herein, we selected two major rainfed rice-growing provinces- Takeo and Prey Veng, based on the information from published reports and government agencies.

As RHEAS has been designed to provide tightly constrained drought and crop yield information in data-scarce regions, the key objective of the study focused on maximizing the use of available scarce observations. Hence, the modeling setup in our study was based on the global calibration of the VIC model by Sheffield and Wood, (2007) and Zhang et al., (2018). Using this calibrated version of VIC, we computed the different hydrologic variables. In addition, the VIC hydrologic model was sufficiently validated against observations. As shown by (Abhishek et al., 2021), the VIC-generated surface soil moisture was validated against the Soil Moisture Active Passive (SMAP: Entekhabi et al., 2015) surface soil moisture (L3_SM_P) data, with the simulated surface soil moisture exhibiting a good agreement (correlation) with SMAP observations (Colliander et al., 2017). As soil moisture from SMAP is regarded as an independent product, validating the model against the observations provides confidence in model stability and performance. Regarding the crop model calibration, the *m*-DSSAT was calibrated by adjusting the crop cultivar genotype coefficients (refer Table. 1). Specifically, the phenological and growth coefficients (see Table 1 for details) were optimized based on an initial set of cultivar genotype coefficients, and yield records for the last 10-12 years. We implemented the Shuffled Complex Evolution (SCE) algorithm to obtain an optimal version of cultivar coefficients by running the crop model multiple times (~350 iterations) (Rahnamay Naeini et al., 2018). As a result, the calibrated crop parameters capture the dynamics of crop growth and development as influenced by the hydrologic parameters and other variables within the system.

At first, the VIC hydrologic model executes nowcast simulations from 1981 through 2011 at a daily time step, computing a large range of hydrologic variables and associated drought indices. These drought indicators are constructed based on the historic climatology (1981-2010) of different hydrologic data streams (e.g., precipitation, soil moisture, etc.). The nowcast simulations are a representation of the current hydrologic states in the region and serve as a base for future prediction. However, as our objective was to examine the efficacy of seasonal yield and drought forecasts for the major growing season (generally the wet season, Jun-Nov), we have devised a comprehensive method of forecasting the yield and hydrologic indicators from 2012 through 2020. For convenience in demonstrating the results, we selected a three-year scenario, each representing a wet (2013), moderate (mod) (2019), and dry (2015) year. As the study focuses on individual administrative units, a comparison of the yield and associated drought conditions for the abovementioned years can be critical in understanding the logic behind the approach.

To test our hypothesis that seasonal climate forecasts improve crop yield predictions over time, we used the multi-model (CFSv2: Climate Forecast System version 2; CCSM4: Community Climate System Model 4.0) ensemble climate forecasts from the NMME seasonal forecasting system as the historic forecast datasets. Each monthly NMME dataset (e.g., Jun) comprises 6-7 months forecast of precipitation and temperature variables. These seasonal forecasts are expressed as probabilities of occurrence below, near, and above normal relative to a long-term climatology of total precipitation and mean temperature. Furthermore, these multi-model ensemble datasets were converted to plausible daily weather sequences (e.g., precipitation, etc.) and ingested as input to the VIC hydrologic model. Using these datasets, the VIC runs forecast simulations for 10 ensemble members to predict the future (Jul-Nov) drought and hydrologic conditions. Following the hydrologic simulation, the *m*-DSSAT crop model subsequently runs a series of (40-50) ensembles to predict the end-of-year (Oct/Nov) crop yield. However, prior to the run, a series of ancillary information, such as cultivar genotype coefficients and agronomic management information (e.g., fertilizer rates), are provided to the crop model. Especially, the type of cultivars, and the amount of fertilization are two key agronomic parameters, that are critical for enhancing (either increasing or decreasing) the crop model performance. Generally, crop models have shown better performance when optimal cultivars and balanced nutrient (fertilizer) application rate information were available. In Cambodia, most of the agricultural setups are rainfed (i.e., mostly paddy rice), thus no irrigation was considered in the study. Although some forms of irrigation activities have recently been introduced around the Tonle Sap Lake,

Table 1
Estimates of genotype coefficients of the rice varieties used in Takeo and Prey Veng provinces of Cambodia. The *m*-DSSAT crop model was calibrated using the Shuffled Complex Evolution (SCE) algorithm for getting an optimized set of phenological and growth parameters.

	Phenological coefficients				Growth Coefficients			
Province	P1	P2R	P5	P2O	G1	G2	G3	G4
Takeo (1)	532.501	122.161	440.195	13.823	68.7939	0.027	1.00	1.20
Prey Veng (2)	554.400	87.7	251.100	13.0	68.67	0.021	1.00	1.15

Phenology genetic coefficients
P1: Time period in °C (above a base temperature of 9 °C) from seedling emergence to the end of juvenile phase. Expressed as growing degree days (GDD)
P2R: Extent to which phase development leading to panicle initiation is delayed for each hour increase in photoperiod above P2O. Expressed as growing degree days (GDD)
P5: Time period from beginning of grain filling to physiological maturity with a base temperature of 9 °C. Unit: GDD
P2O: Critical photoperiod or the longest day length at which development occurs at a maximum rate. Expressed in hours
Growth genetic coefficients
G1: Potential spikelet number coefficients as estimated from the number of spikelets per g of main culm dry weight (less leaf blades and sheaths plus spikes) at anthesis. Unit: Spikelets per g of main culm
G2: Single grain weight under ideal growing conditions, i.e., non-limiting light, water, nutrients, and absence of pests and diseases. Expressed in grams
G3: Tillering coefficient (scalar value) relative to IR64 cultivar under nonlimiting conditions
G4: Temperature tolerance coefficient

herein, we focused on the provinces that have no irrigation.

After the forecast run of drought and crop yield at the end of June, the hydrologic model simulates the present condition (nowcast) for the subsequent month, i.e., July. A similar approach is then followed in the next run, wherein the NMME forecast data of July (i.e., from August to February) is ingested into the RHEAS, and the drought and yield forecast is generated for Aug-Dec. It should be noted that the hydrologic model provides forecast beyond the harvest month (i.e., Nov), however, the crop model is simulated specifically for the long (wet) growing season (i.e., Jun-Nov) as rice is generally harvested in Oct/Nov. Likewise, the process is replicated for each monthly (Jun-Nov) NMME forecast datasets until we reach the end of harvest, i.e., Nov. It is noteworthy that during end of each incremental monthly forecasting of drought and crop yields, the number of months (days) of real forcing data increases. Ultimately, the drought and crop yields simulated for the month of November is done entirely through the real forcing data. A detailed step-by-step flow of the methodology employed is provided in Fig. 4. The general idea behind such an approach is to see the dependence on reducing climate forecasts inputs (i.e., forcings) as we move forward in time during the growing season. For instance, the yield from the Jun NMME forecasts is completely dependent on the historic projection datasets. However, in the Jul NMME forecast run, we see the crop model avails the Jun nowcast and the rest of the NMME forecast to provide the end-of-year yield. Similarly, when we reach Nov, the model is the least dependent on the historic forecast datasets and more on the actual conditions (nowcast).

2.5. Mutual information

Our aim was to examine the relationships between the parameters of

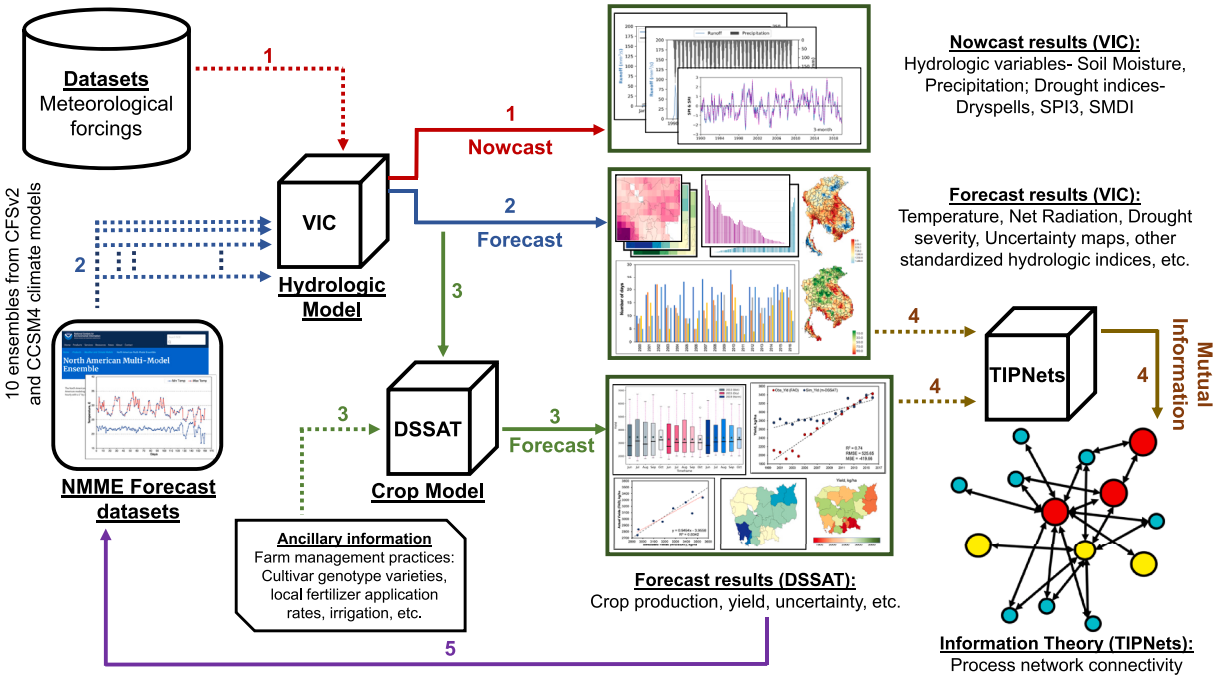

Fig. 4. Detailed representation of the methodology employed within RHEAS for the study. The color (and number) of the arrows represents the steps incorporated in the study chronologically. At the core, the Variable Infiltration Capacity (VIC) hydrologic model loosely couples with the Decision Support System for Agrotechnology Transfer (DSSAT) crop model to provide information on interannual crop yields and associated drought states at current (nowcast) and future (forecast) conditions. For prediction of future states, the forcing datasets consist of multiple coupled models (CFSv2: Climate Forecast System version 2; CCSM4: Community Climate System Model 4) from the North American Multi-Model Ensemble (NMME) forecasting system. Steps 1-4 is carried out for each timeframe (Jun-Oct) over the growing season. Finally, the Temporal Information Partitioning Network (TIPNets) reflects the significant information (or link) shared between the associated variables.

the two coupled dynamic systems, i.e., the hydrologic and crop models, through the information that was exchanged between both systems. As many of these relationships were likely nonlinear, the statistics of IT provide an approach to capture this nonlinearity without requiring any prior assumptions about the variables. Having a general quantification of the connections between different components of the hydrologic (e.g., soil moisture) and crop (e.g., phenology) modeling systems can help identify the inherent parameters/processes affecting interannual crop yields. Specifically, Mutual Information (MI) measures the amount of shared information or synchronization between two variables and is well-equipped to analyze the dynamics of the state variables of a nonlinear coupled system. Here, we treated the hydrologic and drought parameters (e.g., precipitation, dryspells, etc.) as source variables, while the target variable was end-of-year yield. Before MI is computed, we quantify the uncertainties contained in a variable (say $^{i}X^{j}$) using the Shannon entropy (Shannon, 1948):

To measure MI for two variables, X and Y, the uncertainty in a variable ($H(X)$) were computed by:

$$H(X) = -\sum_{i=1}^{N} p(x_i)\ \log_2(p(x_i)) \tag{1}$$

where $H(X)$ (measured in bits) measures the amount of uncertainty contained in X and is based on its probability density function $p(x_i)$ (PDF). Now, using the idea of Shannon entropy, we compute the MI between two variables X (target variable) and Y (source variable). We can understand this easily by considering X as 'crop yield' and Y as 'hydrologic variables or drought parameters'. Thus, the information shared between X and Y can be measured using Eq. 2.

$$MI(X;Y) = H(X) + H(Y) - H(X,Y) = H(X) - H(X|Y) \tag{2}$$

where, MI essentially measures the reduction in uncertainty of X (the target variable) by having this new information about Y (the source variable) and is based on the joint probability of the two variables, and $H(X,Y)$ is the joint entropy. MI can be measured for two variables over a time lag or can be measured without a time lag (often referred as zero-lag MI), provided by:

$$MI(X;Y) = \sum p(x,y)\log_2\left(\frac{p(x,y)}{p(x)p(y)}\right) \tag{3}$$

where, $p(x,y)$ is the joint probability of lagged X and Y variable. To assess the contribution of uncertainty from multiple (source) variables (e.g., precipitation, temperature, etc.) to a single target variable (e.g., crop yield), generally the Normalized mutual information (NMI) is preferred:

$$NMI(X,Y) = \frac{MI(X;Y)}{H(X,Y)} \tag{4}$$

where $NMI(X,Y)$ is the NMI between X and Y. By dividing the mutual information between the source and target variables by the entropy of the target variable, NMI provides a normalized measure that ranges between 0 and 1, wherein a value of 0 indicates no shared information or explanatory power, while a value of 1 indicates that the source variable completely explains the uncertainty in the target variable.

Here, we used a Matlab-based interface, the Temporal Information Partitioning Network (TIPNets: Goodwell and Kumar, 2017) system to quantify the MI between all hydrologic variables and yield. The software first takes the hydrologic time series and yield ensembles as inputs and creates the PDFs of the data using either the fixed interval or kernel density estimation (KDE) binning scheme. KDE is generally employed when there are few data points, but for this work, we had sufficient data to use the fixed interval binning scheme. The joint PDFs of the principal hydrologic parameters, drought indicators, and yield were determined using the fixed interval binning scheme with 10 bins used for computing the PDF (Ruddell and Kumar, 2009a, 2009b). The TIPNets compute the Shannon entropy ($H(X)$) for each variable and MI for all possible variable pairs. Here, we only focused on the zero-lag MI for the relationships with yield, as it was not a time series variable. Statistical significance among the time series variables was found using a shuffled surrogate method where all time series were randomly shuffled and the MI statistics were recomputed using the random series. If the original MI values surpass the random MI values following a t-test at a 95% confidence threshold, then the relationships were considered statistically significant. We performed this analysis for the two study areas over the three years, and the five time periods in each year. The MI values were then normalized by the Shannon entropy of the yield variable to see the portion of yield uncertainty explained by each variable, termed as NMI.

3. Results and discussion

3.1. Efficacy of seasonal climate forecast on crop yield

Our main goal was to understand the evolution of uncertainties of climate forecasts on seasonal crop yields. Based on our working hypothesis, we conducted a monthly forecast of interannual yields based on a 1-month moving window. Herein, we implemented the approach on two major rainfed rice-growing provinces in Cambodia- a) Takeo, and b) Prey Veng for three distinct years- 2013 (wet), 2015 (dry) and 2019 (moderate or mod). These administrative units host large areas of rice production (part of the Mekong Delta) (MRC, 2014) and have been historically impacted by drought/climate extremities, which makes them ideal case studies. To characterize the uncertainties in climate forecast, we used a 40-ensemble yield product from the m-DSSAT crop model for each forecast cycle. Each cycle (represented as I, II, III, IV, V) refers to the nature of input data (Fig. 5), wherein the first cycle (i.e., Jun) of a year comprises 1-month of real meteorologic conditions ('nowcast' that is simulated by the model), while the rest of the input data are based on the ingested NMME seasonal forecast. Similarly, with each subsequent cycle, we have a higher degree of present conditions (nowcast) and lesser reliance on forecasts. Fig. 5a and 5b present the end-of-season yield over Takeo and Prey Veng during the study period. By dividing the analysis into five time intervals, we aimed to capture the temporal dynamics of yield prediction and evaluating their accuracy throughout the growing season. Here, the boxplot with a distinct color for each year shows the extent of simulated yield values for each cycle over a cropping season (i.e., 5 months). For instance, in Takeo, most yield values for Jun 2013 ranged between ~2100 kilogram per hectare (kg/ha) and ~4800 kg/ha (barring the outliers), whereas as we progress in time, the yield values substantially decrease, ranging between ~2500 kg/ha and ~4000 kg/ha (a reduction of ~1200 kg/ha) respectively (Fig. 5a). In line with the observed trend, similar patterns were also observed in other years for both provinces. Specifically, a reduction in yield was recorded, amounting to approximately 400 kg/ha (province Prey Veng, year 2019) (Fig. 5b) and 1300 kg/ha (province Takeo, year 2019) (Fig. 5a). These yield reductions indicate a notable decline in crop productivity during those respective years. Overall, we see a clear indication of reducing uncertainties in yield over the cropping season from planting (i.e., 'I' timeframe or Jun) through harvest (i.e., 'V' timeframe or Oct). In addition, the average yields from the crop model were compared against an independent source of yield records (obtained from Department of Agricultural Land Resources Management, General Directorate of Agriculture, Phnom Penh, Cambodia). In Fig. 5a and 5b, we see the median (simulated) yields (represented by the black line) in accordance with observations (represented by a star mark) over both provinces. When comparing ensembles of simulated yield against observations, the median offers a favorable alternative to the mean due to its resilience to extreme values or outliers. The mean can be significantly distorted by a few outlying values within the ensemble of simulated yields. In contrast, the median, which represents the middle value when the data is arranged in ascending order, diminishes the impact of extreme values, thus yielding a more robust measure of central

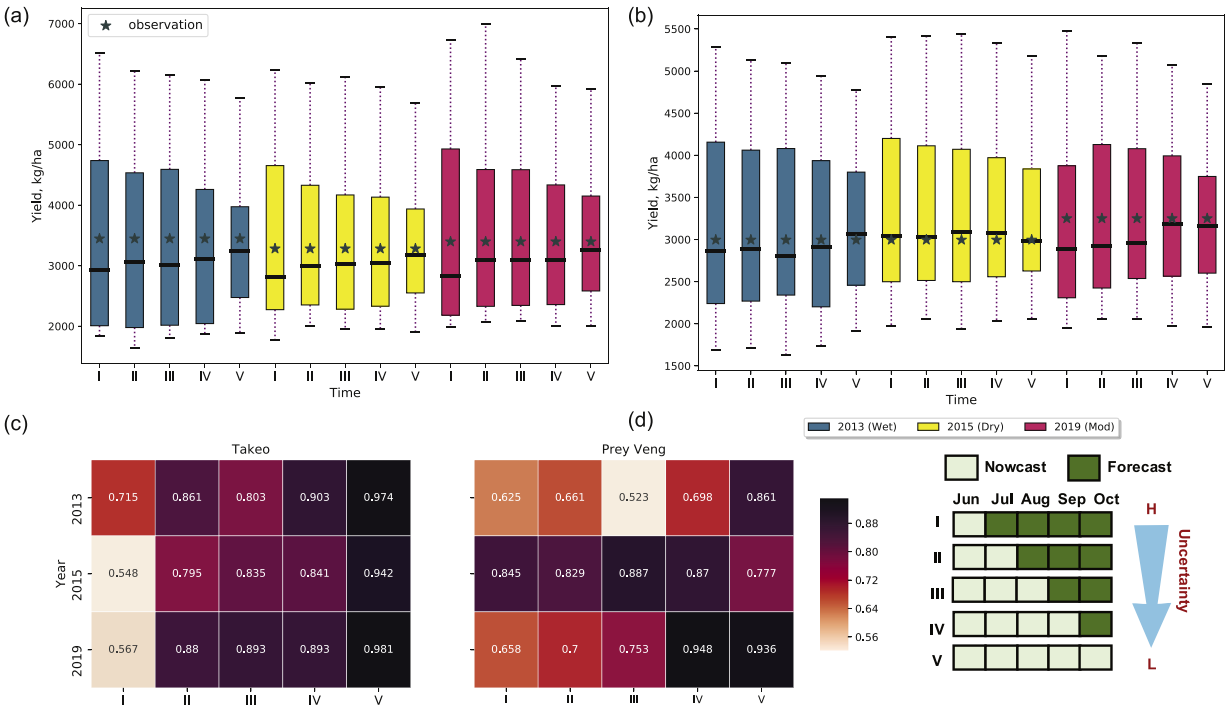

Fig. 5. Uncertainties in yield over the growing season from *m*-DSSAT crop model. a) Boxplot of simulated rice yields from 40 ensembles of the crop model over Takeo during three (wet, dry, mod) distinct years. Here, a season consists of 5 timeframes with a 1-month moving window. Each timeframe represents the end-of-year yield scenarios based on the weightage of nowcast/forecast input forcings. The I, II, III, IV, V timeframes essentially represent the yield values wherein the input forcings comprised of 1-month nowcast and 4-month of forecasts. With each subsequent timeframe, there is an increase in nowcast forcings and decreasing reliance on forecasts. The black line on the boxplots represents the median yield over that particular timeframe, while the grey stars represent the observations. b) Boxplot of simulated yields over Prey Veng during 2013, 2015 and 2019. c-d) Skill of crop yield predictions over the growing season: correlation coefficient between simulated and observed yields over different timeframes in Takeo and Prey Veng.

tendency. Additionally, since observations typically consist of single values, while an ensemble represents multiple values, focusing on the median aligns the comparison by emphasizing the central tendency of the ensemble rather than its mean, thus ensuring a more appropriate and accurate assessment of the model's performance against observations. At the beginning (i.e., 'I' timeframe in 2013 or Jun, 2013), when we have less information on the meteorological conditions, the end-of-season yield (~2930 kg/ha) in Takeo does not reproduce the observations (~3500 kg/ha) (Fig. 5a). However, with subsequent inclusion of actual meteorologic conditions into the model at later timeframes (e.g., IV (Sep) andV (Oct) timeframes), there is an improvement in the yield prediction. Notably, in Takeo, we see the model underestimating yield when compared against observations during I-III timeframes (i.e., Jun-Jul-Aug for all three years, but inches closer towards the actual yields during harvest (V timeframe or Oct). Additionally, the model tends to capture the uncertainty and observed values during the dry (R^2=0.8) and moderate (R^2=0.9) years (as compared to the wet year) (Fig. 5c). Similarly, the model performed considerably better over Prey Veng (than Takeo) in capturing the yield uncertainties whilst accurately predicting/matching the observed yields. While there were noteworthy instances of simulated yield aligning with observations across various timeframes (such as June, July, and October 2015), the model demonstrated a progressive enhancement in its ability to match yield observations in Prey Veng during 2019, achieving an R^2 value of 0.9. Fig. 5c and 5d show the correlation between simulated and observed yields over different timeframes for both provinces. It should be noted that the average simulated yield of Takeo exhibits substantial bias (~300 kg/ha in 2013 and 2015 to ~430 kg/ha in 2019) compared to the observation, especially during the initial timeframes of the year. This large bias can be attributed to the mismatch/absence of: 1) rice cultivar genetic coefficients, and 2) appropriate data on local management practices (e.g., fertilizer rates, planting date, etc.). Although sufficient calibration (of the crop model) was carried out for each province the lack of information on explicit rice cultivars and management practices during the entire period (3 years) could be a major source of uncertainty (see Table. 1 for calibrated values of cultivar genotype coefficients). Additionally, Supplementary Fig. S1 illustrates the relative yield anomaly during different timeframes within a season over both provinces. Hence, having appropriate knowledge about such key information (e.g., agronomic management information, daily weather information, etc.) can be deemed imperative for quantifying yield forecasts. Not only do these results support our hypothesis, but they also provide a substantial ground for using such a unique approach for improving crop yield forecasting.

3.2. *Influence of geophysical and drought variables on crop yield*

To address the uncertainty related to weather, the impact of various hydrologic (and drought) variables were analyzed to examine the contribution of each variable toward the end-of-season yields. We computed the NMI among all the hydrologic and drought variables with simulated rice yields to quantify the synchronous dependence of one variable on another. Normalization of MI is commonly performed to scale its values between 0 and 1, making it more interpretable and comparable across different analyses. Various methods can be employed for this normalization, such as dividing the mutual information by the square root of the product of the entropies of the variables involved. This normalization guarantees that MI falls within the range of [0, 1], irrespective of the variables' specific ranges or units. By normalizing MI, it becomes possible to compare the relationship strengths between variables across diverse time frames or datasets. It is important to note that the range of a variable itself does not directly impact the calculation of normalized mutual information, as the focus lies on the probabilistic distributions of the variables rather than their ranges. However, if a

variable has a limited or constrained range, it may affect its overall variability, potentially influencing its contribution to the mutual information calculation. Here, we focused on crop yield as the target variable to identify the active/predominant parameters impacting yield through the growing season as the forecasts are updated. Fig. 6 shows the NMI between the hydrologic variables and drought indices on crop yields over different timeframes in a year (i.e., how much information is mutually shared between the hydrologic (and drought) variables and crop yield), where the stars highlight the variable with the greatest synchronization with yield. In Takeo, we observed a significant flow of information between the independent variables and crop yield during different timeframes in a season (Fig. 6a). Each timeframe exhibited a varied degree of information flow through a given season, with TMIN and DS exhibiting a consistently higher synchronization with yield over the entire study period, i.e., TMIN had a strong connection with crop yield. During the wet year (2013), TMIN (~0.24 bits/bits, bi/bi), DS (~0.24 bi/bi), and the SMDI (SMDI: (Narasimhan and Srinivasan, 2005) (~0.12 bi/bi) showed higher synchronization, whereas the net Short Wave radiation (SW) (~0.08 bi/bi), and the 3- and 6-month Standardized Precipitation Index (i.e., the SPI3, and SPI6: (McKee et al., 1993) (~0.08 bi/bi) arrayed lower levels of MI. SMDI is a standardized drought index that is based on the surface soil moisture and root zone soil moisture, often used to characterize agricultural drought. It takes into account the soil's water-holding capacity and calculates the moisture deficit by comparing the current soil moisture level with the long-term average. A positive value (e.g., +1.0, +1.5, +2.0) represents wetter conditions (higher than normal long-term average) whereas negative values (e.g., -1.0, -1.5, -2.0) represents drier soil moisture (lower than normal long-term average). Likewise, the 3- and 6-month SPI provides information on the precipitation deficit or surplus over a three-month period respectively. These indices are constructed by accounting the long-term average precipitation for the respective timeframes, enabling the identification of abnormal conditions. The SPI values are an indication of the dryness/wetness in a region, with positive SPI values (i.e., +0.5, +1.0, +1.5, etc.) indicating wetter conditions (i.e., above-average precipitation), and negative SPI values (i.e., -0.5, -1.0, -1.5, etc.) indicating drier conditions (i.e., below-average precipitation). Having knowledge about these parameters offers valuable insights into drought conditions, helping stakeholders respond effectively. Additionally, we saw an increase in the information flow from the start of the cropping season (Jun) until harvest (Oct) during the wet year. This essentially means that the strength between the nodes, each representing a geophysical parameter (e.g., DS, SW, etc.), increased over time. In 2013, there was an increase of ~47% (DS) and ~128% (TMIN), while a few parameters saw a decrease of ~18% (PREC) and 24% (SW) during the same period. A similar pattern was observed during the dry year (2015), wherein the information flow was higher during the later timeframes (i.e., Sep, and Oct). These percentages represent the relative increase or decrease in the values of DS and SW from the beginning (June) to the end (October) of the season. The fluctuations observed throughout the year, with variables increasing and then decreasing, are captured by these percentage changes, which highlights the seasonal dynamics and variations in DS and SW. Although the mod year (2019) depicted irregular behavior, it showed a higher synchronization than did the dry year. Overall, we observed yield exhibiting a higher synchronization with the hydrologic variables (10 out of 15 months) compared to the drought indicators with TMIN as the strongest link. Also, the drought variables displayed strong links during the earlier months (Jun, Jul) of the wet and dry years, while the latter months (Sep, Oct) array a strong synchronization with the hydrologic variables. During the initial months of both wet and dry years, we found strong connections between drought indices (e.g., Standardized Precipitation Index and Standardized Soil Moisture Deficit Index) and crop yield. This indicates that drought conditions significantly influenced crop performance in the early stages of the growing season. These findings were consistent with previous research (Hussain et al., 2016) emphasizing the importance of

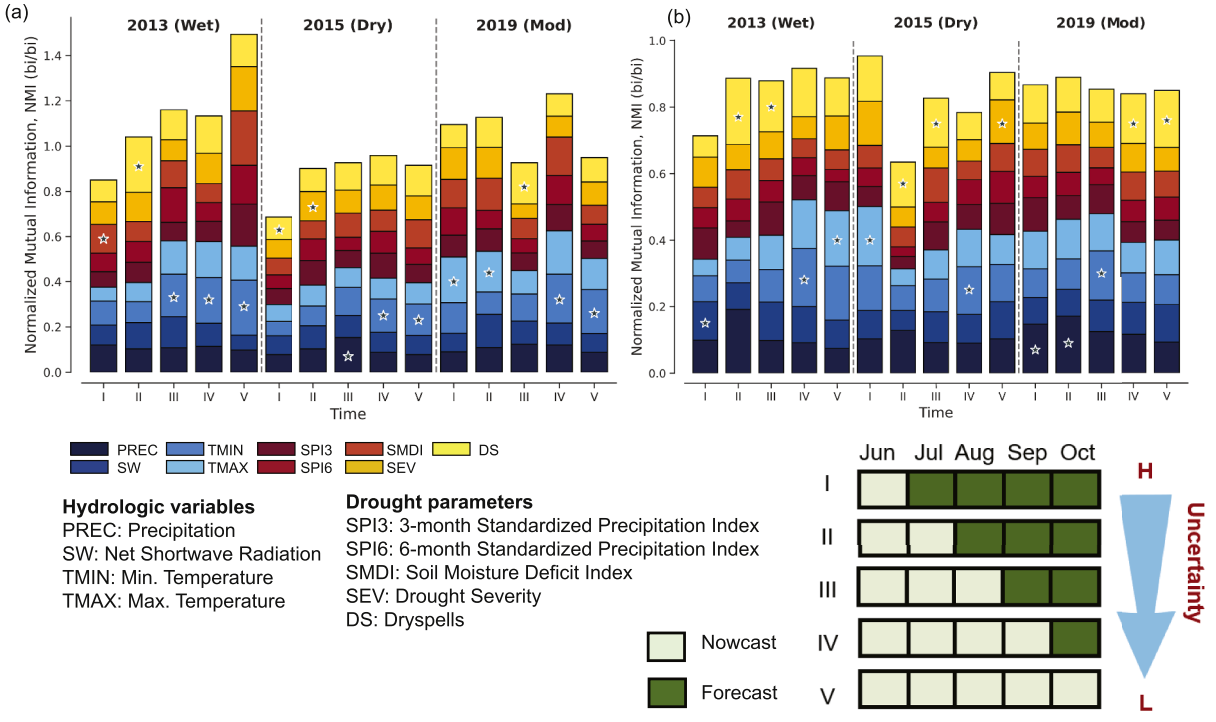

Fig. 6. Stacked bar plot showing the NMI between different hydrologic variables (PREC, TMIN, TMAX, SW) and crop (rice) yield, and NMI between drought indicators (SPI3, SMDI, SPI6, SEV, DS) and crop yield: a) Mutual exchange of information over Takeo; b) Mutual exchange of information over Prey Veng. The stacked bars represent the contribution of the independent variables and crop yield at different timeframes. Shades of blue color (navy to sky blue) represents different hydrologic variables while the shades of red color (red to yellow) represent the different drought indicators. The star (in each stacked bar) represents the variable/indicator having the maximum influence on crop yield during that time window.

early-season drought on subsequent crop outcomes. However, as we approached the harvest period, we noticed a shift in the significance of variables. Hydrologic factors like precipitation and soil moisture exhibited stronger associations with crop yields compared to the drought parameters. This suggests that hydrologic variables played a more influential role in determining final crop yields as the growing season progressed. This shift can be attributed to the direct impact of hydrologic variables on water availability and nutrient supply during critical growth and development stages. The increasing synchronization between crop yield and hydrologic variables in the later months of the growing season can be explained by crop physiology. As crops mature and approach the harvest stage, their water requirements and sensitivity to soil moisture conditions become more pronounced. Consequently, hydrologic variables that directly influence soil moisture and water availability play a critical role in determining crop yields during this crucial period. Likewise, we saw a similar pattern of information flow in Prey Veng wherein DS, TMIN, TMAX, and PREC had a substantial impact on yield (Fig. 6b). Compared to Takeo, the amount of information mutually exchanged between the models was relatively stable throughout the months, which could potentially be due to the better performance of the crop model over Prey Veng. In contrast to Takeo which arrayed strong predominant links with hydrologic variables (10 out of 15 months), the NMI for Prey Veng showcased strong linkages with 8 hydrologic and 7 drought variables over the study period, with DS dominating the information sharing with yield compared to other variables (6 out of 15 months). It is also noteworthy that a dominant link between two nodes at an initial timeframe may not necessarily have a significant synchronization at a later timeframe. For instance, PREC exhibited the strongest synchronization with crop yields at the start of the year in 2019, but the strength between the links gradually decreased (or became insignificant) during the later timeframe (i.e., IV and V timeframe in 2019). Instead, DS maintained a consistent link throughout the season and had the highest information flow during the end-of-season timeframe, thus making it the most suitable variable that considerably influences crop yield.

3.3. Process network connectivity

In our study, we examined the various network connections between the hydrologic variables, drought indicators, and end-of-season yield in Takeo and Prey Veng region. To achieve this, we constructed a network plot that provides a comprehensive visualization of the relationships between hydrologic variables and crop yield during different time frames (see Fig. 7 and Fig. 8). We visualized these connections using process networks by analyzing the strength and direction of the relationships among these variables throughout the growing season. Typically, it involves creating a network plot, where each variable is represented as a node (or point), with each node encompassing a particular variable under consideration (e.g., PREC, TMAX, TMIN, DS, etc.) (Fig. 7 and Fig. 8). These nodes typically represent the timeseries (based on the PDF) of a particular variable. These nodes are connected via arrows that illustrates the connection between different variables.

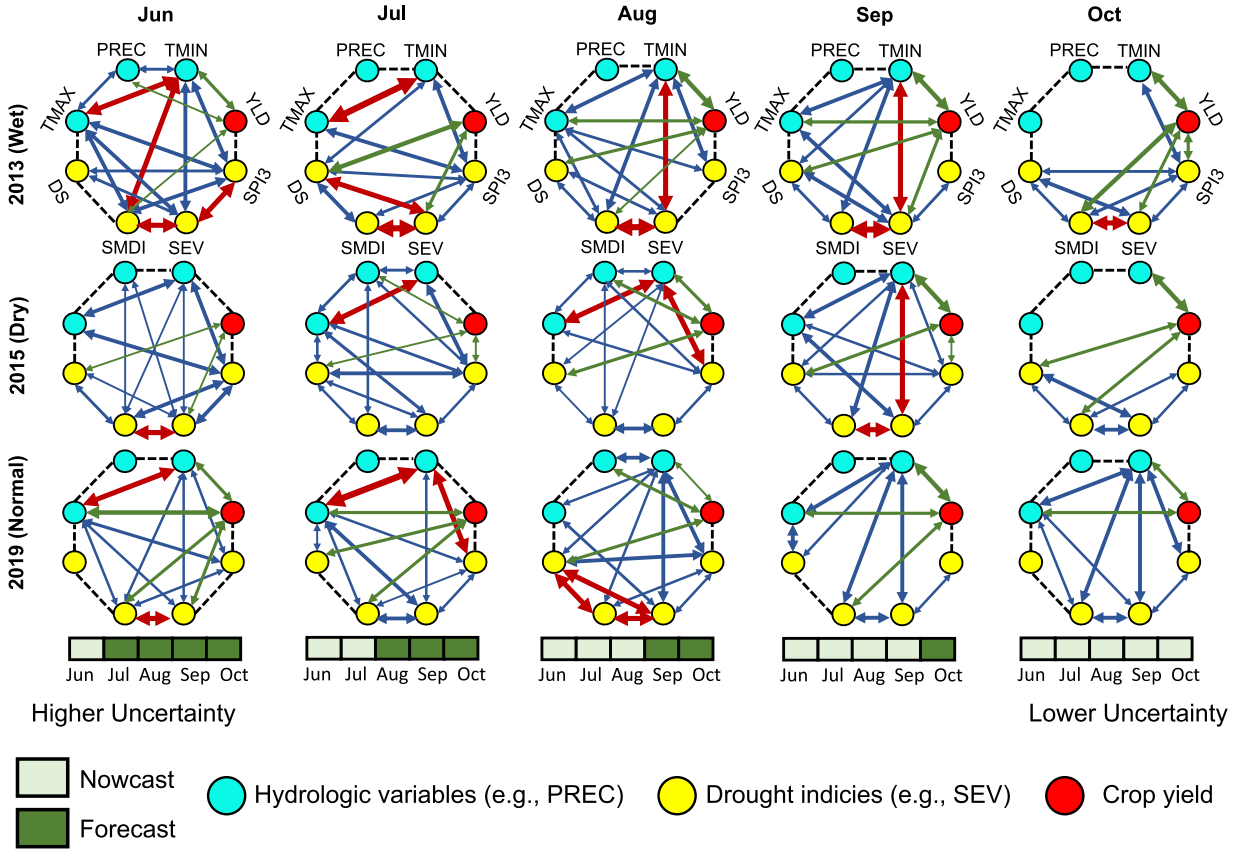

Fig. 7. Process network connectivity. Establishing the potential links (connectivity) between the core hydrologic variables (represented by aqua colored circles), associated drought indicators (represented by yellow circles) and end-of-season crop (rice) yields (represented by red circles) over the growing season (Jun-Oct) in Takeo during the wet year (2013) (top panel), dry year (2015) (middle panel), and mod year (2019) (bottom panel) respectively. Each variable, also called a node, represents the timeseries of the related variable. The thickness of the links represents the strength of the relationship. The color variations in the plot reflect the varying degrees of significance, with darker colors indicating stronger mutual information. Red arrows represent the most dominant network connections, while the blue arrows represent moderate-low strength, and green arrows represent the strength of the independent variables that have significant influence on seasonal yields. The dashed black line merely provides completeness and visual aesthetics to the network plot. TMIN: Min. Temperature; TMAX: Max. Temperature; PREC: Precipitation; SPI3: 3-month Standardized Precipitation Index; SEV: Drought Severity; SMDI: Soil Moisture Deficit Index; DS: Dryspells; YLD: Crop yield.

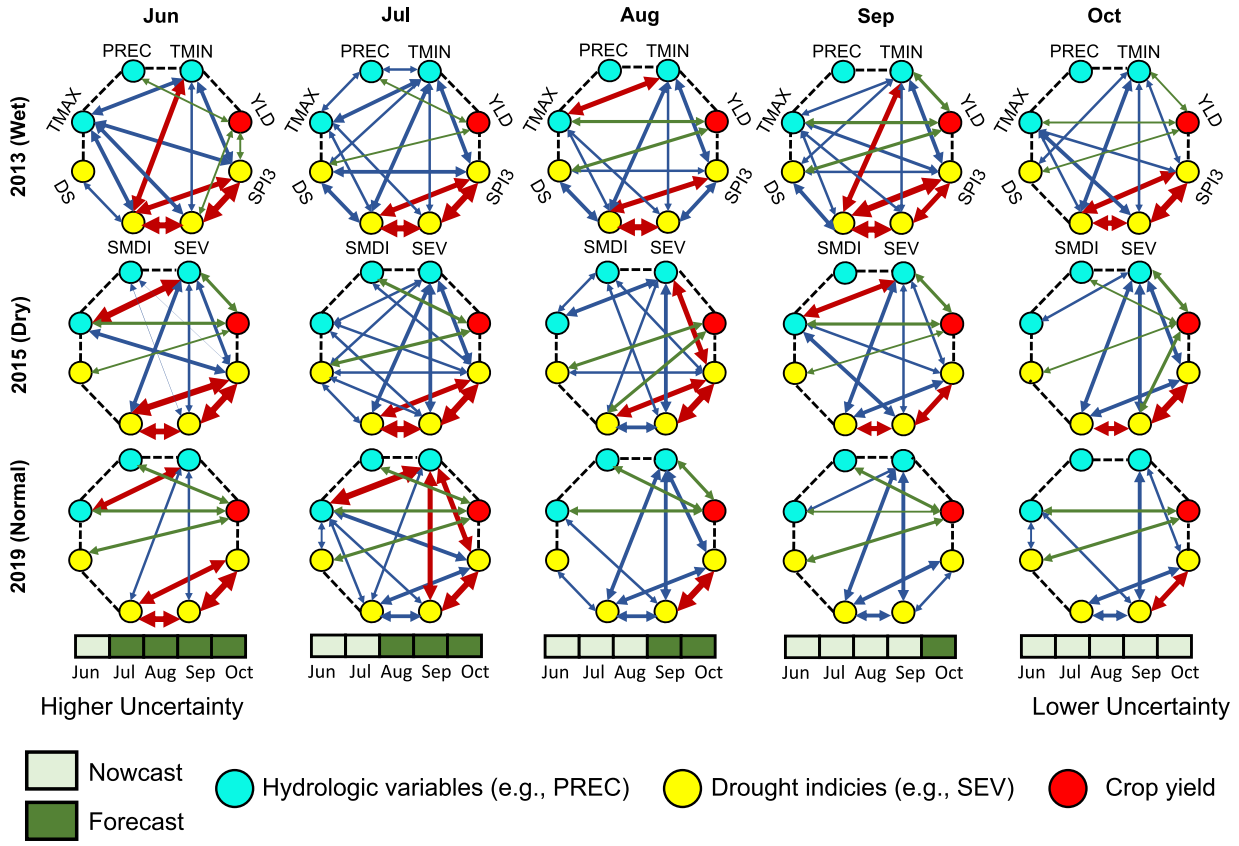

Fig. 8. Network connectivity over Prey Veng. Process networks showing the connectivity between the hydrologic parameters (aqua circles), drought indicators (yellow circles) and yield (red circles) over Prey Veng during a wet (2013) (top panel), dry (2015) (middle panel) and normal (2019) (bottom panel) year respectively. TMIN: Min. Temperature; TMAX: Max. Temperature; PREC: Precipitation; SPI3: 3-month Standardized Precipitation Index; SEV: Drought Severity; SMDI: Soil Moisture Deficit Index; DS: Dryspells; YLD: Crop yield.

These connections between different variables are established based on the NMI values calculated by the TIPNets model. In simple terms, the arrows show the statistical dependence between two variables, taking into account their joint distribution and individual distributions. Additionally, the intensity and thickness of the arrows indicate the strength of the connection between variables, and help in visualizing the dependencies within the system. For e.g., a thinner arrow indicates a weaker connection, signifying a low NMI value. Conversely, a thicker arrow represents a stronger association, indicating a high NMI value. The varying thicknesses of the arrows are a visual depiction of the various relationships between different variables/parameters, aiding in the identification of dominant interactions. By examining the network plots as a whole, we identified the influential nodes (or variables) that act as key drivers within the system during different time period.

To provide more clarity on the strength of the connections, we categorized the thickness and color of the arrows based on the NMI values. Specifically, the red and blue color arrows were employed to identify distinct relationships between the hydrologic variables and drought parameters. On the other hand, the green color arrows were used to depict the connections between hydrologic variables and crop yield, as well as drought parameters and crop yield. Similarly, the thickness of the arrows represented the strength and intensity of the connections between two variables based on the NMI values. A thicker red arrow (for e.g., SPI3 and SEV in Jun, 2013) represents a strong connection, indicating NMI values greater than 0.55 bi/bi. Similarly, a relevantly thin red arrow depicts strong connection with NMI values ranging between 0.4-0.54 bi/bi. Likewise, the blue arrows represent the moderate connections that have NMI values ranging between 0.3-0.4 bi/bi, while thin blue arrows show the weak connections with NMI values ranging between 0.1 to 0.3 bi/bi. Classifying the connections based on the NMI values allows for a better understanding of the varying strength of the connections between the variables. Regarding the black dashed line in the network plots, they were merely used to ensure the visual completeness and aesthetics of the network plot, and do not carry any specific meaning or significance.

Due to limited space within the network plot, we prioritized variables based on their importance, relevance, and existing research findings. Hence, we illustrated the process connectivity between selected variables—SPI3, SMDI, SEV, DS, TMAX, TMIN, PREC and crop yield (represented as YLD), thus leading to the exclusion of other variables such as SPI6 and SW from the network plot. During the initial months of the study years (specifically, Jun and Jul), we observed significant network connections between variables, particularly in the wet year in Takeo (see Fig. 7). In Jun, there was a strong association between SEV and SMDI (with NMI of approximately 0.55 bi/bi), while in Jul, a strong connection was observed between SPI3 and SPI6 (with NMI of approximately 0.6-0.7 bi/bi). Conversely, the links between DS and TMAX (NMI of ~0.08 bi/bi), and PREC and SEV (NMI of ~0.06 bi/bi) were weak or insignificant. However, as we progressed further in time (e.g., Sep, Oct), the connections between the variables became more explicit and prominent, i.e., the variables that exhibited higher NMI values in the previous months (i.e., Jun and Jul) continued to show strong connections, while the other links with lower NMI values during that period (i.e., Jun and Jul) became weak or insignificant. Overall, in 2013, DS and SPI3, SMDI and SPI3, DS and SEV, TMIN and SPI3 represented a consistent association with each other during the growing season. This outcome was seen as a result of the changing nature of input forcings onto the model. To elaborate, by incorporating real forcing data (i.e., less observed weather data from NMME and more input data sampled from climatology for the remaining season), the bulk of the noisy links

weakened over time and certain strong links persisted throughout the growing season. For e.g., in 2013, we consistently observed strong connections between SPI3 and SMDI, SPI6 and SPI3, SEV and SPI3, and SMDI and DS, while other links became insignificant over the course of the growing season. This information not only indicates the robustness and stability of the dominant connections, but enables us to identify key parameters that aids in understanding the influence of different variables on water availability. Additionally, we observed significant connections during the dry and mod years. Specifically, during the dry year (2015), we found a consistent relationship between SMDI and SEV, SPi3 and SEV, and DS and SMDI, while in 2019, SEV and TMIN, SMDI, and TMIN, SMDI and SEV, TMAX and TMIN, and TMIN and SPI3 exhibited a strong consistent relation. Thus, having information on the relationships between the key hydrologic variables and drought parameters during different time periods enables us to determine alternative agricultural outcomes.

Similarly, we analyzed the NMI to quantify the influence of the independent variables (i.e., both hydrologic variables and drought parameters) on crop yields. The variables that had the greatest influence (i.e., variables exhibiting higher NMI values) on crop yields were shown in the network plot by green arrows. During the wet year, we found TMIN, SMDI, SEV, and SPI3 had the maximum influence on end-of-year yields, indicating a significant statistical relationship with crop yield. Similarly, in the dry year (2015), we observed a significant statistical relationship between crop yield and DS, TMIN, while TMAX, DS, and TMIN exhibited a significant relationship during the mod year (2019). Likewise, we conducted a similar process network connectivity analysis over Prey Veng (see Fig. 8), and the findings aligned with those of Takeo. The analysis revealed a similar conclusion—1) Connection density is higher when there is more reliance on seasonal forecasts (i.e., during the initial months, Jun-Jul), 2) Over time the information flow networks become distinct and stronger as real (nowcast) data is incorporated into the model, and the consistent links persist, and 3) TMIN, TMAX and DS showed the maximum influence on end-of-season yield, reinforcing the significance of these variables in determining crop outcomes. Thus, the interpretation of such connections can substantially help us understand the variability and predictability of crop yields during different time in a season.

4. Summary and conclusion

We employed a novel technique for evaluating the efficacy of historic seasonal climate forecasts based on a multi-model ensemble. Based on the current application, RHEAS performed reasonably well in quantifying the yield uncertainties associated with seasonal climate forecasts. Particularly, the *m*-DSSAT crop model was able to capture the uncertainties in end-of-season yield, whilst significantly complementing observations over the growing season. To characterize the uncertainties, the *m*-DSSAT model provides a 40-ensemble yield scenario based on a 1-month timeframe moving window. These model ensembles simulate yield based on the nature of: 1) meteorological forcings (from the VIC hydrologic model), 2) cultivar varieties, 3) fertilizer application rates, whilst capturing the model uncertainties, either arising from model parameters or agronomic practices. The results from the study showed the extent of yield values becoming better constrained (i.e., reduction of uncertainty) with each subsequent timeframe over both provinces. Overall, the model provided a near-accurate yield forecast with reasonable bias (that improves over every forecast/month) at longer lead times (see Fig. 5). In addition, the model performed reasonably well in matching the observations over both provinces. However, due to the absence of real input forcings during the initial months (Jun-Jul-Aug) of the season, there were substantial discrepancies (gaps or bias) between the simulated and observed yields. However, with the availability of more nowcasts in subsequent months (Sep-Oct), the average simulated yield mimics the observations, thus significantly improving the initial bias. As no parameters were changed other than the meteorologic forcings, this inherently connotes the quality of initial forcings as one of the critical drivers for yield prediction. Not only do these results support our hypothesis, but also provide a substantial ground for using such an approach to improve crop yield forecasting for agriculture sustainability.

Even though the impacts of weather-related uncertainties can be modeled satisfactorily, understanding the relationships between the key hydrologic and agriculture variables is essential for identifying the dominant pathways and variables affecting interannual crop yields. Herein, we employed the concept of MI to discern the synchronous flow of information between the key hydrologic and agriculture variables. Particularly, we examined the information mutually exchanged between the hydrologic (and drought) parameters (independent variable) and crop yields (dependent variable). Our results showed a significant increase in NMI during the initial timeframes (Jun-Jul) as compared to the later timeframes (Sep-Oct), most notably during the wet (Takeo, Prey Veng) and dry year (Takeo) (see Fig. 6). In other words, such behavior essentially implies the significant transfer of information at a timeframe wherein we have the maximum weightage of the present (nowcast) conditions, i.e., when our yield predictions are completely based on real simulated data (and not on historic forecasts). Whilst higher yield uncertainties arrayed less NMI during the Jun-Jul timeframe, there were few instances of relatively stable (2019, Prey Veng) or irregular (2019, Takeo and 2015, Prey Veng) behavior of the independent variables towards yield. Over the study duration, DS and TMIN maintained a consistent level of statistical significance with yield, i.e., TMIN and DS were the dominant parameters having the maximum influence on interannual yields. In addition, we observed that hydrologic variables, specifically temperature, demonstrated a stronger correlation with yield in 10 out of 15 months in the Takeo region (i.e., temperature have a greater impact on crop yield in Takeo). Conversely, in Prey Veng, we found a higher correlation between yield and 8 hydrologic variables and 7 drought variables throughout the study period (i.e., a broader range of hydrologic and drought-related factors influence crop yield in Prey Veng). It is noteworthy, while it suggests a higher significance of particular variables during a certain timeframe, it is important to consider that the significance may vary depending on specific years, regions, and prevailing conditions.

Additionally, we looked at the process network connectivity between the key hydrologic variables, drought indicators, and yield. Our results indicated the prevalence of multiple network connections between the independent variables during the start of the season (Jun) (see Fig. 7 and Fig. 8). However, the bulk of these links diminish over time and the strong prominent links persist at the end of the season. Such a pattern can be attributed to the forcing information provided to the model, i.e., the associated uncertainties in the forecast. During harvest, the model simulates the hydrologic parameters based on the real conditions (nowcast) which reduce the noisy links and highlight the explicit networks. Likewise, we presented the dominant links between the independent variables and yield, with DS and TMIN being the most influential parameters over both provinces. Thus, to obtain a sound and reliable forecast of crop (rice) yield with manageable uncertainties for a growing season, it is desirable to have a seasonal climate forecast that encapsulate the information on DS and TMIN optimally. Currently, all the available seasonal forecasts do not have such information. Hence, including skillful information on DS md TMIN will effectively enhance the forecast capability with uncertainty estimates essential for agricultural sustainability.

Code availability

Additional details about the RHEAS framework and source code can be obtained at https://gitlab.com/kandread/RHEAS. The process network connectivity approach was based on the TIPNets software, available at https://github.com/HydroComplexity/TIPNet.

CRediT authorship contribution statement

Abhijeet Abhishek: Investigation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Mantha S. Phanikumar:** Formal analysis, Writing – review & editing. **Alicia Sendrowski:** Methodology, Formal analysis, Resources. **Konstantinos M. Andreadis:** Conceptualization, Methodology, Writing – review & editing. **Mahya G.Z. Hashemi:** Formal analysis, Data curation. **Susantha Jayasinghe:** Resources, Data curation, Validation. **P.V. Vara Prasad:** Resources, Data curation, Validation. **Roberts J. Brent:** Resources, Data curation. **Narendra N. Das:** Conceptualization, Methodology, Resources, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Narendra Narayan Das reports financial support was provided by NASA.

Data availability

Data will be made available on request.

Acknowledgements

This study was funded by the joint initiative between the National Aeronautics and Space Administration (NASA) and the United States Agency for International Development (USAID), under the SERVIR-Mekong project (Grant no. ROSES #80NSSC21K0403). The SERVIR-Mekong is housed at the Asian Disaster Preparedness Center (ADPC) in Bangkok, Thailand, the primary implementer of the RHEAS integrated framework for addressing critical challenges in food security, agriculture, and climate-risk disasters. We thank the support from the SERVIR Coordination Office, especially Dr. Ashutosh Limaye, and the NASA Program Manager Dr. Nancy Searby for their guidance and facilitating the project collaboration with USAID. We also acknowledge the support of Dr. Prakash Jha who provided the historic crop yield data over different provinces of Cambodia.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.agrformet.2023.109683.

References

Abhishek, A., Das, N.N., Ines, A.V.M., Andreadis, K.M., Jayasinghe, S., Granger, S., Ellenburg, W.L., Dutta, R., Hanh Quyen, N., Markert, A.M., Mishra, V., Phanikumar, M.S., 2021. Evaluating the impacts of drought on rice productivity over Cambodia in the Lower Mekong Basin. J Hydrol (Amst) 599, 126291. https://doi.org/10.1016/j.jhydrol.2021.126291.

Ahmadalipour, A., Moradkhani, H., Castelletti, A., Magliocca, N., 2019. Future drought risk in Africa: Integrating vulnerability, climate change, and population growth. Science of the Total Environment 662, 672–686. https://doi.org/10.1016/j.scitotenv.2019.01.278.

Alizadeh, M.R., Adamowski, J., Nikoo, M.R., Aghakouchak, A., Dennison, P., Sadegh, M., 2020. A century of observations reveals increasing likelihood of continental-scale compound dry-hot extremes. Sci. Adv.

Andreadis, K.M., Das, N., Stampoulis, D., Ines, A., Fisher, J.B., Granger, S., Kawata, J., Han, E., Behrangi, A., 2017. The regional hydrologic extremes assessment system: A software framework for hydrologic modeling and data assimilation. PLoS One 12, 1–22. https://doi.org/10.1371/journal.pone.0176506.

Basso, B., Liu, L., 2019. Seasonal crop yield forecast: Methods, applications, and accuracies, 1st ed, Advances in Agronomy. Elsevier Inc. https://doi.org/10.1016/bs.agron.2018.11.002.

Becker, E., van den Dool, H., 2016. Probabilistic seasonal forecasts in the North American Multimodel Ensemble: A baseline skill assessment. J Clim 29, 3015–3026. https://doi.org/10.1175/JCLI-D-14-00862.1.

Brown, J.N., Hochman, Z., Holzworth, D., Horan, H., 2018. Seasonal climate forecasts provide more definitive and accurate crop yield predictions. Agric For Meteorol 260–261, 247–254. https://doi.org/10.1016/j.agrformet.2018.06.001.

Challinor, A., 2011. Forecasting food. Nat Clim Chang 1, 103–104. https://doi.org/10.1038/nclimate1098.

Challinor, A.J., Koehler, A.K., Ramirez-Villegas, J., Whitfield, S., Das, B., 2016. Current warming will reduce yields unless maize breeding and seed systems adapt immediately. Nat Clim Chang 6, 954–958. https://doi.org/10.1038/nclimate3061.

Clarke, B., Otto, F., Stuart-Smith, R., Harrington, L., 2022. Extreme weather impacts of climate change: an attribution perspective. Environmental Research: Climate 1, 012001. https://doi.org/10.1088/2752-5295/ac6e7d.

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

Dai, A., 2013. Increasing drought under global warming in observations and models. Nat Clim Chang 3, 52–58. https://doi.org/10.1038/nclimate1633.

Eini, M.R., Javadi, S., Delavar, M., Monteiro, J.A.F., Darand, M., 2019. High accuracy of precipitation reanalyses resulted in good river discharge simulations in a semi-arid basin. Ecol Eng 131, 107–119. https://doi.org/10.1016/j.ecoleng.2019.03.005.

Entekhabi, B.D., Njoku, E.G., Neill, P.E.O., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., Kimball, J., Piepmeier, J.R., Koster, R.D., Martin, N., Mcdonald, K.C., Moghaddam, M., Moran, S., Reichle, R., Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L., Zyl, J. Van, 2015. (SMAP) Mission 98.

Ermida, S.L., DaCamara, C.C., Trigo, I.F., Pires, A.C., Ghent, D., Remedios, J., 2017. Modelling directional effects on remotely sensed land surface temperature. Remote Sens Environ 190, 56–69. https://doi.org/10.1016/j.rse.2016.12.008.

Feng, P., Wang, B., Liu, D.L., Waters, C., Xiao, D., Shi, L., Yu, Q., 2020. Dynamic wheat yield forecasts are improved by a hybrid approach using a biophysical model and machine learning technique. Agric For Meteorol 285–286. https://doi.org/10.1016/j.agrformet.2020.107922.

Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sens Environ 114, 168–182. https://doi.org/10.1016/j.rse.2009.08.016.

Fujimori, S., Hasegawa, T., Krey, V., Riahi, K., Bertram, C., Bodirsky, B.L., Bosetti, V., Callen, J., Després, J., Doelman, J., Drouet, L., Emmerling, J., Frank, S., Fricko, O., Havlik, P., Humpenöder, F., Koopman, J.F.L., van Meijl, H., Ochi, Y., Popp, A., Schmitz, A., Takahashi, K., van Vuuren, D., 2019. A multi-model assessment of food security implications of climate change mitigation. Nat Sustain 2, 386–396. https://doi.org/10.1038/s41893-019-0286-2.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci Data 2, 150066. https://doi.org/10.1038/sdata.2015.66.

Gampe, D., Zscheischler, J., Reichstein, M., O'Sullivan, M., Smith, W.K., Sitch, S., Buermann, W., 2021. Increasing impact of warm droughts on northern ecosystem productivity over recent decades. Nat Clim Chang 11, 772–779. https://doi.org/10.1038/s41558-021-01112-8.

Goodwell, A.E., Jiang, P., Ruddell, B.L., Kumar, P., 2020. Debates—Does Information Theory Provide a New Paradigm for Earth Science? Causality, Interaction, and Feedback. Water Resour Res 56, 1–12. https://doi.org/10.1029/2019WR024940.

Goodwell, A.E., Kumar, P., 2017. Temporal Information Partitioning Networks (TIPNets): A process network approach to infer ecohydrologic shifts. Water Resource Research 53, 5899–5919. https://doi.org/10.1002/2016WR020216.

Guo, H., Bao, A., Liu, T., Ndayisaba, F., He, D., Kurban, A., De Maeyer, P., 2017. Meteorological Drought Analysis in the Lower Mekong Basin Using Satellite-Based Long-Term CHIRPS Product. Sustainability 9, 901. https://doi.org/10.3390/su9060901.

Hagemann, S., Dümenil Gates, L., 2001. Validation of the hydrological cycle ECMWF and NCEP reanalyses using the MPI hydrological discharge model. Journal of Geophysical Research Atmospheres 106, 1503–1510. https://doi.org/10.1029/2000jd900568.

Hamman, J.J., Nijssen, B., Bohn, T.J., Gergel, D.R., Mao, Y., 2018. The variable infiltration capacity model version 5 (VIC-5): Infrastructure improvements for new applications and reproducibility. Geosci Model Dev 11, 3481–3496. https://doi.org/10.5194/gmd-11-3481-2018.

Han, E., Ines, A.V.M., Koo, J., 2019. Development of a 10-km resolution global soil profile dataset for crop modeling applications. Environmental Modelling and Software 119, 70–83. https://doi.org/10.1016/j.envsoft.2019.05.012.

Hansen, J, Challinor, A, Ines, A, Wheeler, T, Moron, V., Hansen, Jw, Challinor, A, Ines, A, Wheeler, T, Moron, V., Hansen, J.W., Challinor, Andrew, Ines, Amor, Wheeler, Tim, Moron, V., 2006. Translating climate forecasts into agricultural terms: advances and challenges To cite this version: HAL Id: hal-02894588 Translating climate forecasts into agricultural terms: advances and challenges.

Hoang, L.P., van Vliet, M.T.H., Kummu, M., Lauri, H., Koponen, J., Supit, I., Leemans, R., Kabat, P., Ludwig, F., 2019. The Mekong's future flows under multiple drivers: How climate change, hydropower developments and irrigation expansions drive

hydrological changes. Science of The Total Environment 649, 601–609. https://doi.org/10.1016/j.scitotenv.2018.08.160.
Holzworth, D.P., Snow, V., Janssen, S., Athanasiadis, I.N., Donatelli, M., Hoogenboom, G., White, J.W., Thorburn, P., 2015. Agricultural production systems modelling and software: Current status and future prospects. Environmental Modelling and Software 72, 276–286. https://doi.org/10.1016/j.envsoft.2014.12.013.
Hussain, M., Waqas-ul-Haq, M., Farooq, S., Jabran, K., Farroq, M., 2016. The impact of seed priming and row spacing on the productivity of different cultivars of irrigated wheat under early season drought. Experimental Agriculture 52 (3), 477–490.
Ines, A.V.M., Das, N.N., Hansen, J.W., Njoku, E.G., 2013. Assimilation of remotely sensed soil moisture and vegetation with a crop simulation model for maize yield prediction. Remote Sens Environ 138, 149–164. https://doi.org/10.1016/j.rse.2013.07.018.
Innes, P.J., Tan, D.K.Y., Van Ogtrop, F., Amthor, J.S., 2015. Effects of high-temperature episodes on wheat yields in New South Wales. Australia. Agric For Meteorol 208, 95–107. https://doi.org/10.1016/j.agrformet.2015.03.018.
Jägermeyr, J., Müller, C., Ruane, A.C., Elliott, J., Balkovic, J., Castillo, O., Faye, B., Foster, I., Folberth, C., Franke, J.A., Fuchs, K., Guarin, J.R., Heinke, J., Hoogenboom, G., Iizumi, T., Jain, A.K., Kelly, D., Khabarov, N., Lange, S., Lin, T.S., Liu, W., Mialyk, O., Minoli, S., Moyer, E.J., Okada, M., Phillips, M., Porter, C., Rabin, S.S., Scheer, C., Schneider, J.M., Schyns, J.F., Skalsky, R., Smerald, A., Stella, T., Stephens, H., Webber, H., Zabel, F., Rosenzweig, C., 2021. Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. Nat Food 2, 873–885. https://doi.org/10.1038/s43016-021-00400-y.
Jha, B., Kumar, A., Hu, Z.Z., 2019. An update on the estimate of predictability of seasonal mean atmospheric variability using North American Multi-Model Ensemble. Clim Dyn 53, 7397–7409. https://doi.org/10.1007/s00382-016-3217-1.
Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. European Journal of Agronomy 18, 235–265. https://doi.org/10.1016/S1161-0301(02)00107-7.
Kadiyala, M.D.M., Jones, J.W., Mylavarapu, R.S., Li, Y.C., Reddy, M.D., 2015. Identifying irrigation and nitrogen best management practices for aerobic rice-maize cropping system for semi-arid tropics using CERES-rice and maize models. Agric Water Manag 149, 23–32. https://doi.org/10.1016/j.agwat.2014.10.019.
Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Jenne, R., Joseph, D., 1996. The NCEP/NCAR 40-Year Reanalysis Project. Bull Am Meteorol Soc 77, 437–471. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.
Kang, H., Sridhar, V., Mainuddin, M., Trung, L.D., 2021. Future rice farming threatened by drought in the Lower Mekong Basin. Sci Rep 11, 1–15. https://doi.org/10.1038/s41598-021-88405-2.
Karthikeyan, L., Chawla, I., Mishra, A.K., 2020. A review of remote sensing applications in agriculture for food security: Crop growth and yield, irrigation, and crop losses. J Hydrol (Amst) 586, 124905. https://doi.org/10.1016/j.jhydrol.2020.124905.
Katsanos, D., Retalis, A., Michaelides, S., 2016. Validation of a high-resolution precipitation database (CHIRPS) over Cyprus for a 30-year period. Atmos Res 169, 459–464. https://doi.org/10.1016/j.atmosres.2015.05.015.
Konapala, G., Mishra, A.K., Wada, Y., Mann, M.E., 2020. Climate change will affect global water availability through compounding changes in seasonal precipitation and evaporation. Nat Commun 11, 1–10. https://doi.org/10.1038/s41467-020-16757-w.
Konduri, V.S., Kumar, J., Hargrove, W.W., Hoffman, F.M., Ganguly, A.R., 2020. Mapping crops within the growing season across the United States. Remote Sens Environ 251, 112048. https://doi.org/10.1016/j.rse.2020.112048.
Kushwaha, N.L., Rajput, J., Shirsath, P.B., Sena, D.R., Mani, I., 2022. Seasonal climate forecasts (SCFs) based risk management strategies: A case study of rainfed rice cultivation in India. Journal of Agrometeorology 24, 10–17. https://doi.org/10.54386/jam.v24i1.775.
Lacasa, J., Messina, C.D., Ciampitti, I., 2023. A probabilistic framework for forecasting maize crop yield response to agricultural inputs with sub-seasonal climate predictions. Environmental Research Letters. https://doi.org/10.1088/1748-9326/acd8d1.
Lal, R., Stewart, B.A. (Eds.), 2018. Soil and Climate, 1st ed. CRC Press. https://doi.org/10.1201/b21225.
Lauri, H., Räsänen, T.A., Kummu, M., 2014. Using Reanalysis and Remotely Sensed Temperature and Precipitation Data for Hydrological Modeling in Monsoon Climate: Mekong River Case Study. J Hydrometeorol 15, 1532–1545. https://doi.org/10.1175/jhm-d-13-084.1.
Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. J Geophys Res 99, 14415. https://doi.org/10.1029/94JD00483.
Liu, S., Lu, P., Liu, D., Jin, P., 2007. Pinpointing source of Mekong and measuring its length through analysis of satellite imagery and field investigations. Geo-spatial Information Science 10, 51–56. https://doi.org/10.1007/s11806-007-0011-6.
Mainuddin, M., Kirby, M., Hoanh, C.T., 2013. Impact of climate change on rainfed rice and options for adaptation in the lower Mekong Basin. Natural Hazards 66, 905–938. https://doi.org/10.1007/s11069-012-0526-5.
Maurer, E.P., Hidalgo, H.G., Das, T., Dettinger, M.D., Cayan, D.R., 2010. The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. Hydrol Earth Syst Sci 14, 1125–1138. https://doi.org/10.5194/hess-14-1125-2010.
Mavromatis, T., 2015. Crop–climate relationships of cereals in Greece and the impacts of recent climate trends. Theor Appl Climatol 120, 417–432. https://doi.org/10.1007/s00704-014-1179-y.
McKee, T.B., Doesken, N.J., Kleist, J., 1993. The relationship of drought frequency and duration to time scales. Eighth Conference on Applied Climatology. https://doi.org/10.1002/jso.23002.
Minoli, S., Jägermeyr, J., Asseng, S., Urfels, A., Müller, C., 2022. Global crop yields can be lifted by timely adaptation of growing periods to climate change. Nat Commun 13. https://doi.org/10.1038/s41467-022-34411-5.
MRC, 2014. Crop production for food security and rural poverty Baseline and pilot modelling.
Müller, C., Franke, J., Jägermeyr, J., Ruane, A.C., Elliott, J., Moyer, E., Heinke, J., Falloon, P.D., Folberth, C., Francois, L., Hank, T., Izaurralde, R.C., Jacquemin, I., Liu, W., Olin, S., Pugh, T.A.M., Williams, K., Zabel, F., 2021. Exploring uncertainties in global crop yield projections in a large ensemble of crop models and CMIP5 and CMIP6 climate scenarios. Environmental Research Letters 16. https://doi.org/10.1088/1748-9326/abd8fc.
Narasimhan, B., Srinivasan, R., 2005. Development and evaluation of Soil Moisture Deficit Index (SMDI) and Evapotranspiration Deficit Index (ETDI) for agricultural drought monitoring. Agric For Meteorol 133, 69–88. https://doi.org/10.1016/j.agrformet.2005.07.012.
Ortiz-Bobea, A., Ault, T.R., Carrillo, C.M., Chambers, R.G., Lobell, D.B., 2021. Anthropogenic climate change has slowed global agricultural productivity growth. Nat Clim Chang 11, 306–312. https://doi.org/10.1038/s41558-021-01000-1.
Padrón, R.S., Gudmundsson, L., Decharme, B., Ducharne, A., Lawrence, D.M., Mao, J., Peano, D., Krinner, G., Kim, H., Seneviratne, S.I., 2020. Observed changes in dry-season water availability attributed to human-induced climate change. Nat Geosci 13, 477–481. https://doi.org/10.1038/s41561-020-0594-1.
Paudel, D., de Wit, A., Boogaard, H., Marcos, D., Osinga, S., Athanasiadis, I.N., 2023. Interpretability of deep learning models for crop yield forecasting. Comput Electron Agric 206. https://doi.org/10.1016/j.compag.2023.107663.
Rahnamay Naeini, M., Yang, T., Sadegh, M., AghaKouchak, A., Hsu, K.lin, Sorooshian, S., Duan, Q., Lei, X., 2018. Shuffled Complex-Self Adaptive Hybrid EvoLution (SC-SAHEL) optimization framework. Environmental Modelling and Software 104, 215–235. https://doi.org/10.1016/j.envsoft.2018.03.019.
Ray, D.K., Gerber, J.S., Macdonald, G.K., West, P.C., 2015. Climate variation explains a third of global crop yield variability. Nat Commun 6. https://doi.org/10.1038/ncomms6989.
Rifai, S.W., Li, S., Malhi, Y., 2019. Coupling of El Niño events and long-term warming leads to pervasive climate extremes in the terrestrial tropics. Environmental Research Letters 14. https://doi.org/10.1088/1748-9326/ab402f.
Ruddell, B.L., Kumar, P., 2009a. Ecohydrologic process networks: 1. Identification. Water Resour Res 45, 1–23. https://doi.org/10.1029/2008WR007279.
Ruddell, B.L., Kumar, P., 2009b. Ecohydrologic process networks: 2. Analysis and characterization. Water Resour Res 45, 1–14. https://doi.org/10.1029/2008WR007280.
Sendrowski, A., Passalacqua, P., 2017. Process connectivity in a naturally prograding river delta. Water Resour Res 53, 1841–1863. https://doi.org/10.1002/2016WR019768.
Shannon, C.E., 1948. A Mathematical Theory of Communication. Bell System Technical Journal 27, 623–656. https://doi.org/10.1002/j.1538-7305.1948.tb00917.x.
Sheffield, J., Wood, E.F., 2007. Characteristics of global and regional drought, 1950–2000: Analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle. J Geophys Res 112, D17115. https://doi.org/10.1029/2006JD008288.
Slater, L.J., Villarini, G., Bradley, A.A., 2019. Evaluation of the skill of North-American Multi-Model Ensemble (NMME) Global Climate Models in predicting average and extreme precipitation and temperature over the continental USA. Clim Dyn 53, 7381–7396. https://doi.org/10.1007/s00382-016-3286-1.
Son, N.T., Chen, C.F., Chen, C.R., Chang, L.Y., Minh, V.Q., 2012. Monitoring agricultural drought in the Lower Mekong Basin using MODIS NDVI and land surface temperature data. International Journal of Applied Earth Observation and Geoinformation 18, 417–427. https://doi.org/10.1016/j.jag.2012.03.014.
Shukla, P.R., Skea, J., Slade, R., Diemen, R. van, Haughey, E., Malley, J., Pathak, M., Pereira, J.P., 2019. Foreword Technical and Preface. Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems 35–74.
Subash, N., Ram Mohan, H.S., 2012. Evaluation of the impact of climatic trends and variability in rice-wheat system productivity using Cropping System Model DSSAT over the Indo-Gangetic Plains of India. Agric For Meteorol 164, 71–81. https://doi.org/10.1016/j.agrformet.2012.05.008.
Thiesen, S., Darscheid, P., Ehret, U., 2019. Identifying rainfall-runoff events in discharge time series: A data-driven method based on information theory. Hydrol Earth Syst Sci 23, 1015–1034. https://doi.org/10.5194/hess-23-1015-2019.
Thilakarathne, M., Sridhar, V., 2017. Characterization of future drought conditions in the Lower Mekong River Basin. Weather Clim Extrem 17, 47–58. https://doi.org/10.1016/j.wace.2017.07.004.
Thober, S., Kumar, R., Sheffield, J., Mai, J., Schäfer, D., Samaniego, L., 2015. Seasonal soil moisture drought prediction over Europe using the North American Multi-Model Ensemble (NMME). J Hydrometeorol 16, 2329–2344. https://doi.org/10.1175/JHM-D-15-0053.1.
Tippett, M.K., Ranganathan, M., L'Heureux, M., Barnston, A.G., DelSole, T., 2019. Assessing probabilistic predictions of ENSO phase and intensity from the North American Multimodel Ensemble. Clim Dyn 53, 7497–7518. https://doi.org/10.1007/s00382-017-3721-y.

Togliatti, K., Archontoulis, S.V., Dietzel, R., Puntel, L., VanLoocke, A., 2017. How does inclusion of weather forecasting impact in-season crop model predictions? Field Crops Res 214, 261–272. https://doi.org/10.1016/j.fcr.2017.09.008.

Toté, C., Patricio, D., Boogaard, H., van der Wijngaart, R., Tarnavsky, E., Funk, C., 2015. Evaluation of satellite rainfall estimates for drought and flood monitoring in Mozambique. Remote Sens (Basel) 7, 1758–1776. https://doi.org/10.3390/rs70201758.

Vogel, E., Donat, M.G., Alexander, L.V., Meinshausen, M., Ray, D.K., Karoly, D., Meinshausen, N., Frieler, K., 2019. The effects of climate extremes on global agricultural yields. Environmental Research Letters 14. https://doi.org/10.1088/1748-9326/ab154b.

Wang, L., Ren, H.L., Xu, X., Huang, B., Wu, J., Liu, J., 2022. Seasonal-Interannual Predictions of Summer Precipitation Over the Tibetan Plateau in North American Multimodel Ensemble. Geophys Res Lett 49. https://doi.org/10.1029/2022GL100294.

Zhang, Y., Pan, M., Sheffield, J., Siemann, A.L., Fisher, C.K., Liang, M., Beck, H.E., Wanders, N., MacCracken, R.F., Houser, P.R., Zhou, T., Lettenmaier, D.P., Pinker, R. T., Bytheway, J., Kummerow, C.D., Wood, E.F., 2018. A Climate Data Record (CDR) for the global terrestrial water budget: 1984–2010. Hydrol Earth Syst Sci 22, 241–263. https://doi.org/10.5194/hess-22-241-2018.

Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D.B., Huang, Y., Huang, M., Yao, Y., Bassu, S., Ciais, P., Durand, J.L., Elliott, J., Ewert, F., Janssens, I.A., Li, T., Lin, E., Liu, Q., Martre, P., Müller, C., Peng, S., Peñuelas, J., Ruane, A.C., Wallach, D., Wang, T., Wu, D., Liu, Z., Zhu, Y., Zhu, Z., Asseng, S., 2017. Temperature increase reduces global yields of major crops in four independent estimates. Proc Natl Acad Sci U S A 114, 9326–9331. https://doi.org/10.1073/pnas.1701762114.