AGU ADVANCING EARTH AND SPACE SCIENCES

JAMES | Journal of Advances in Modeling Earth Systems®

RESEARCH ARTICLE
10.1029/2023MS004023

Multi-Decadal Soil Moisture and Crop Yield Variability—A Case Study With the Community Land Model (CLM5)

Theresa Boas[1,2,3], **Heye Bogena**[1], **Dongryeol Ryu**[3], **Andrew Western**[3], and **Harrie-Jan Hendricks Franssen**[1,2]

[1]Forschungszentrum Jülich, Institute of Bio- and Geosciences: Agrosphere (IBG-3), Jülich, Germany, [2]Centre for High-Performance Scientific Computing in Terrestrial Systems: HPSC TerrSys, Jülich, Germany, [3]Department of Infrastructure Engineering, University of Melbourne, Parkville, VIC, Australia

Key Points:

- Land surface models (LSMs) with integrated crop models can be used to quantify the impact of climate change on agro-ecosystems
- The potential value of LSMs for agricultural purposes depends on their ability to adequately simulate inter-annual variability of yield
- The representation of plant hydraulics and the soil moisture regime play key role in accurately simulating agro-ecosystems

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

T. Boas,
t.boas@fz-juelich.de

Citation:

Boas, T., Bogena, H., Ryu, D., Western, A., & Hendricks Franssen, H.-J. (2024). Multi-decadal soil moisture and crop yield variability—A case study with the Community land model (CLM5). *Journal of Advances in Modeling Earth Systems*, *16*, e2023MS004023. https://doi.org/10.1029/2023MS004023

Received 13 SEP 2023
Accepted 20 AUG 2024

Author Contributions:

Conceptualization: Theresa Boas, Heye Bogena, Dongryeol Ryu, Andrew Western, Harrie-Jan Hendricks Franssen
Data curation: Theresa Boas
Formal analysis: Theresa Boas
Funding acquisition: Harrie-Jan Hendricks Franssen
Investigation: Theresa Boas
Methodology: Theresa Boas, Heye Bogena, Dongryeol Ryu,

Abstract While the impacts of climate change on global food security have been studied extensively, the capability of emerging tools that couple land surface processes and crop growth in reproducing inter-annual yield variability at regional scale remains to be tested rigorously. In this study, we analyzed the effects of weather variations between years (1999–2019) on regional crop productivity for two agriculturally managed regions with contrasting climate and cropping conditions: the German state of North Rhine-Westphalia (DE-NRW) and the Australian state of Victoria (AUS-VIC), using the latest version of the Community Land Model (CLM5) and the WFDE5 (WATCH Forcing Data methodology applied to ECMWF reanalysis version 5) reanalysis. Overall, the simulation results were able to reproduce the total annual crop yields of certain crops, while also capturing the differences in total yield magnitudes between the domains. However, the simulations showed limitations in correctly capturing inter-annual differences of crop yield compared to official yield records, which resulted in relatively low correlation coefficients between 0.07 and 0.39 in AUS-VIC and between 0.11 and 0.42 in DE-NRW. The mean absolute deviation of simulated winter wheat yields was up to 4.6 times lower compared to state-wide records from 1999 to 2019. Our results suggest the following limitations of CLM5: (a) limitations in simulating yield responses from plant hydraulic stress; (b) errors in simulating soil moisture contents compared to satellite-derived data; and (c) errors in the representation of cropland in general, for example, crop parameterizations and human influences.

Plain Language Summary This study evaluates how year-to-year weather variations impact crop yield predictions for two regions, North Rhine-Westphalia in Germany and Victoria in Australia changes. We use the community land model (CLM5) land surface model in combination with reanalysis weather data to investigate the model performance with respect to the representation of crop phenology, plant water stress, and soil moisture. Our results showcase the model's ability to predict total annual crop yield magnitudes for both regions, while also capturing the differences between the respective simulation domains. However, year-to-year changes in crop yield were lower in simulation results compared to official records, which indicated a lack of model sensitivity toward drought stress and general limitations in the representation of agricultural land. This research systematically assesses CLM5 model performance over arable land and provides useful insights into limitations of CLM5 that can help guide future empirical and technical model improvements.

1. Introduction

Agricultural production and management are closely connected to weather and climate conditions and farming yields are significantly affected by inter-annual weather variability. In addition to changes in annual average temperatures and shifts in seasonality, recent climate projections also indicate an increasing number of extreme weather events and a higher intensity of such events, which poses a new challenge for agriculture (Challinor et al., 2014; Deryng et al., 2014; Levis et al., 2018; Rosenzweig et al., 2014; Tai et al., 2014; Urban et al., 2012). The impacts of climate change on food security and agricultural land are a research topic with high relevance to society. In addition, the fluxes of water, energy and carbon associated with agriculture (use of irrigation and fertilizer, timing of crop growth and fallow periods, etc.) can have implications for local and regional weather and climate, and biochemistry (Sacks et al., 2009).

Numerical modeling of Earth system components plays a vital role in assessing the impacts of climate change, exploring adaptation strategies and their impact on various parts of the terrestrial system. Land surface models (LSMs) such as the Community Land Model (CLM) are essential tools for studying changes in response to

Andrew Western, Harrie-Jan Hendricks Franssen
Project administration: Harrie-Jan Hendricks Franssen
Resources: Theresa Boas
Software: Theresa Boas
Supervision: Heye Bogena, Dongryeol Ryu, Andrew Western, Harrie-Jan Hendricks Franssen
Validation: Theresa Boas
Visualization: Theresa Boas
Writing – original draft: Theresa Boas
Writing – review & editing: Theresa Boas, Heye Bogena, Dongryeol Ryu, Andrew Western, Harrie-Jan Hendricks Franssen

weather conditions and are particularly valuable for examining the effects of climate change on agricultural land at larger spatial scales. Prognostic simulations of land surface models (LSMs) and global crop models can be used to quantify the impact of climate change on agro-ecosystems and study the response of agricultural land to inter-annual weather variations. While both contribute to our understanding of Earth's systems, LSMs encompass a broader focus on land surface processes, including natural ecosystems, while global crop models specialize in simulating and analyzing agricultural systems at a global scale. For example, results from the Global Gridded Crop Model Intercomparison (GGCMI; Franke et al., 2020; Jägermeyr et al., 2021) offer valuable insights into long-term productivity trends and the adaptive capacity of the agricultural system under different climate scenarios on a global scale and have been applied to investigate various questions, for example, on challenges for food production, future crop yields and irrigation water demand (e.g., Blanchard et al., 2017; Jägermeyr et al., 2021; Müller et al., 2015; Wada et al., 2013). The Agricultural Model Intercomparison and Improvement Project (AgMIP) is another important example of a research initiative focused on improving agricultural models and enhancing our understanding of climate change impacts on food security and developing integrated assessment tools for decision-making in agriculture (e.g., Asseng et al., 2019; Cammarano et al., 2020; Kimball et al., 2019; Rosenzweig et al., 2013, 2014; Tumbo et al., 2020; White et al., 2013). Multiple studies have showcased assessments of AgMIP data sets at the regional scale and for a variety of crops, which revealed significant variations in climate change impacts on wheat yields across different regions, emphasizing the need for tailored adaptation strategies (e.g., Asseng et al., 2019; Cammarano et al., 2020; Kimball et al., 2019; Rosenzweig et al., 2014).

The impacts of climate change on food security have received considerable attention in recent years. Still, the consequences of altered weather patterns on yearly yield variations remains an important area of interest. The potential value of LSMs for these purposes largely depends on their ability to adequately simulate the crop productivity variability, which has not been tested rigorously for dryland and dominantly rain-fed cropping regions, where yield is vulnerable to drought and heat stresses. The inter-annual variability of model output is an important performance measure for land surface models (LSMs) as it reflects the ability of the model to capture the natural variability observed in the real world over multiple years. Assessing and understanding inter-annual variability of terrestrial fluxes is crucial for various applications, including climate projections, agricultural management, water resources management, and ecosystem dynamics. Reliable predictions of regional crop yield variability can help to design agricultural adaptation and mitigation strategies and can provide useful information for local stakeholders and policy makers.

In order to adequately represent inter-annual variability of crop growth, it is crucial for the model to sufficiently represent the soil moisture regime and corresponding vegetative drought stress in response to changes in precipitation amounts, specifically in dominantly rain-fed areas. While an increasing number of studies focus on incorporating irrigation and human water use in LSMs and hydrological models in general (e.g., McDermid et al., 2023; Pokhrel et al., 2012, 2016; Shah et al., 2019; Xia et al., 2022; Yassin et al., 2019), many challenges still remain in representing rain-water limited agricultural regimes in general circulation models and LSMs.

The realistic representation of many processes is still a challenge in LSMs. Major challenges arise from the complex representation of soil-plant-atmosphere feedbacks, root water uptake and plant responses to environmental stress, land use and land cover change, spatial heterogeneity in soil properties, vegetation cover and topography, and uncertainties in model parameterizations (Blyth et al., 2021; Dagon et al., 2020; Fisher & Koven, 2020; Franks et al., 2018; Huntzinger et al., 2013; Lombardozzi et al., 2020; Sabot et al., 2022; Sulis et al., 2019; Trugman et al., 2018). The ability of LSMs to capture the impacts of land use change on biogeochemistry and hydrology is often limited by the oversimplification of human influences on land use and land cover, such as agricultural practices and management decisions. Additionally, multiple studies found that the representation of plant responses to environmental stress need to be improved in global LSMs (De Kauwe, Zhou, et al., 2015; Franks et al., 2018; Sabot et al., 2022; Sulis et al., 2019). In a recent study, Boas et al. (2023) highlighted the potential value of combining CLM5 with seasonal weather forecasts. They also identified limitations of the model in capturing inter-annual variations in agricultural characteristics over a short period of four cropping seasons.

Focusing on soil water, root water uptake plays an important role in rain-fed agricultural systems and is often simplified in LSMs, which can affect simulated vegetation growth, productivity and water use efficiency of the plants (De Kauwe, Zhou, et al., 2015; Sulis et al., 2019). Most LSMs, including earlier versions of CLM, utilize

soil moisture stress parameterizations where water stress is based on a plant wilting factor (calculated with the soil water matric potential values corresponding to plant dependent parameters for fully open and fully closed stomata conditions) (Lawrence et al., 2019). In contrast to this physical representation, the active role of roots in redistributing water within the soil profile has been investigated by numerous studies and different formulations of root hydraulic redistribution have been included in LSMs, highlighting the relevance of atmospheric processes and carbon and nutrient cycling (e.g., Li et al., 2012; Ryel et al., 2002; Sulis et al., 2019; Tang et al., 2015; Yan & Dickinson, 2014; Zheng & Wang, 2007). CLM5 includes a plant hydraulic stress model that replaced the empirical soil moisture stress formulation from earlier model versions (Lawrence et al., 2019). This plant hydraulic stress scheme simulates vegetation water potential for every segment in the soil-root-stem-leaf system and includes a stress formulation where leaf water potential is used to attenuate photosynthesis. This plant hydraulic framework provides a better physical basis for multiple processes represented in CLM, such as the attenuation of photosynthesis and transpiration during drought conditions (Lawrence et al., 2019). A similar approach has been presented in Sulis et al. (2019), where a macroscopic root water uptake model was introduced in CLM (version 4.0) that also explicitly simulated the leaf water potential at stomatal closure defining water stress conditions for the plants. They found that root hydraulic properties control transpiration during dry periods and that the roots distribution induced a larger variability in the hydraulic model response (Sulis et al., 2019). Their modified model resulted in a good correlation of simulated and observed transpiration fluxes for a winter wheat test site and a more distinct response under water stress conditions compared to default model simulations (Sulis et al., 2019). The subject of plant response to drought stress also encompasses the representation of stomatal conductance. In a recent study by Sabot et al. (2022), multiple empirical and optimization formulations for stomatal conductance were evaluated using a simplified LSM framework. They found that the selection of the stomatal conductance model could considerably influence the simulation of carbon and water exchange in global models. Further factors that can impact inter-annual variations in yield variability and that also need to be considered in the numerical representation of the system are changes in crop management practices in relation to technical advances, public policies, and farming techniques such as fertilization or double-cropping, as well as pests, diseases, and floods (Lombardozzi et al., 2020).

Furthermore, the accuracy of LSMs also heavily depends on their complex parameterizations that aim to account for a wide range of variability in soil and vegetation types. Uncertainties arise due to the limited availability of observational data for parameter estimation and validation as well as the complexity of the parameterizations themselves, leading to potential errors in model predictions (Boas et al., 2021; Huntzinger et al., 2013; Lombardozzi et al., 2020; Lu et al., 2017; Sulis et al., 2015). In addition to improvements in model parameterization (e.g., through new methods for parameter estimation and uncertainty quantification), the study by Fisher and Koven (2020) highlights the role of data assimilation techniques, including the use of remote sensing and machine learning, that can help improve the accuracy of LSM predictions (e.g., Dagon et al., 2020; Pinnington et al., 2020, 2021).

This study aims to analyze the model performance across two regions with different climates and dominated by rain-fed agriculture, for a period spanning multiple decades. Specifically, we assessed the ability of the CLM5 with its prognostic crop module to capture the inter-annual variability of crop yield, soil moisture and plant water stress over two simulation domains: one covering large parts of the south-east Australian wheat belt in the state of Victoria in Australia (AUS-VIC) and the other extending over the state of North Rhine-Westphalia in Germany (DE-NRW). Simulations were conducted over the two regional domains, which are in different climate zones, and forced with the global bias-adjusted reanalysis data set WFDE5 (Cucchi et al., 2020). We compared our simulation results with recorded yields and examined which variables (i.e., seasonal rainfall, root zone soil moisture) dominantly drive changes in CLM5-predicted total yield and yield variability. Additionally, the simulated multi-decadal near-surface soil moisture was compared with two reference data sets, the combined ESA-CCI product (Dorigo et al., 2017) and the satellite-derived SMAP L3 soil moisture product (Entekhabi et al., 2016). In this study, we provide an overview on multi-decadal model performance across two model domains, both of which are dominated by rain-fed agriculture and with state-wide agricultural yield statistics available as validation data. The results of this follow-up study provide valuable insights for the use of CLM5 in agricultural landscapes. We addressed both the model's skill and limitations in predicting long-term variations in annual crop productivity and soil moisture levels at the regional scale, which allowed us to highlight specific areas in need for further investigation.

2. Methodology

2.1. Land Surface Model

In this study, all simulation were carried out with CLM5, which is the first model version that includes a fully prognostic crop module (Lawrence et al., 2018, 2019; Lombardozzi et al., 2020). In CLM5, human management is represented by fertilization and irrigation. Other aspects of agricultural management, such as residue management and soil tillage, or ecological factors affecting crops, such as pests, diseases, or wildlife damage, are also not accounted for in CLM5 at this stage of development. Crop growth and phenology are simulated based on atmospheric factors (i.e., incoming shortwave and longwave radiation, atmospheric pressure, relative humidity, wind speed and temperature) and water availability from irrigation and precipitation (soil moisture). Besides, crop biomass and yield depend on nutrient availability in the soil. Fertilization is represented in a simplified scheme by adding prescribed amounts of nitrogen directly to the soil mineral pool. The possibility of simulating crop growth and development in CLM5 enables a broader and more accurate approach to address economic challenges and questions in land use change and agriculture (e.g., Lobell et al., 2006). Furthermore, a new plant hydraulic stress formulation was introduced in CLM5 that explicitly simulates the transport of water within the soil-root-leaf system, as well as the plant-mediated vertical hydraulic redistribution of soil water from wet to dry soil layers was implemented (Lawrence et al., 2018, 2019). While CLM5 does not explicitly model groundwater dynamics, it indirectly considers its effects on crop growth through its representation of soil moisture dynamics (redistribution within the soil column) and groundwater discharge and recharge. Groundwater is simulated with explicit representation of the saturated and unsaturated zone, using soil thickness and impermeable bedrock as a zero-flux boundary. Soil profile depths are based on a spatially explicit soil thickness data product by Pelletier et al. (2016). Additionally, the stomatal conductance scheme was updated in CLM5 to the approach proposed by Medlyn et al. (2011). The new model formulations led to better performance in simulating ecosystem water fluxes, vegetation water stress, and productivity, thus providing a basis for an improved plant water use and water stress simulation in future applications of the model (Lawrence et al., 2019).

The plant hydraulic stress routine simulates water transport in the soil-root-stem-leaf system by explicitly calculating water potential gradients based on Darcy's Law for porous media flow (Lawrence et al., 2018, 2019 and references therein). The representation of either positive or negative soil-to-root fluxes depending on the water potential gradients allows for a plant-mediated vertical hydraulic redistribution of soil water from wet to dry soil layers through vegetation tissue (Lawrence et al., 2018, 2019).

Water potential gradients in the soil-root-stem-leaf system (water fluxes from soil to root, from root to stem and in between the plant segments) are modeled at each time step as follows (Lawrence et al., 2018, 2019):

$$q = kA\left(\psi_1 - \psi_2\right) \tag{1}$$

where q is the flux of water spanning the segment between ψ_1 and ψ_2 [mmH_2O/s], ψ_1–ψ_2 is the gradient in water potential across the segments [mmH_2O], k is the hydraulic conductance [s^{-1}] and A is the area basis [m^2/m^2].

The segment's resistance to hydraulic stress is calculated with a sigmoidal curve function, where hydraulic conductance decreases as water potentials decrease. The maximum segment conductance is multiplied by a sigmoidal function that accounts for the percentage loss of conductivity using the water potential at 50% loss of conductivity (p_{50}) and a shape parameter (Lawrence et al., 2018):

$$k = k_{max} \cdot 2^{-\left(\frac{\psi_1}{p50}\right)^{c_k}} \tag{2}$$

where k_{max} is the maximum segment conductance [s^{-1}], p_{50} is the water potential at 50% loss of conductivity [mmH_2O], ψ_1 is the water potential of the lower segment terminus [mmH_2O] and c_k is the vulnerability curve shape-fitting parameter [-]. Parameters such as k_{max}, p_{50} and c_k strongly control the modeled plant hydraulic stress routine and thus determine the capability of the plant to extract water from the soil and to resist hydraulic stress. These routines are physically constrained by the plant hydraulic parameterization after Kennedy et al. (2019) which until now contains the same parameters for all crops.

In this routine, the vegetation water potential responds to water supply and transpiration demand (i.e., plant water demand), and transpiration demand is dependent on stomatal conductance. The leaf stomatal conductance and leaf photosynthesis are modeled for sunlit and shaded leaves separately based on the approaches after Medlyn et al. (2011), and Farquhar et al. (1980) for C_3 plants and Collatz et al. (1992) for C_4 plants (Lawrence et al., 2018). Adapted from Medlyn et al. (2011), the leaf stomatal resistance is calculated using the net leaf photosynthesis, the vapor pressure deficit and the CO_2 concentration at the leaf surface with plant-specific slope parameters based on de Kauwe et al. (2015a) and Lin et al. (2015) as follows (Lawrence et al., 2018):

$$\frac{1}{r_s} = g_s = g_o + 1.6\left(1 + \frac{g_1}{\sqrt{D}}\right)\frac{A_n}{\frac{c_s}{P_{atm}}} \tag{3}$$

where r_s is leaf stomatal resistance [s m^2/μmol], g_o is the minimum stomatal conductance [μmol/m^2/s], A_n is leaf net photosynthesis [μmolCO_2/m^2/s], g_1 is the plant dependent slope parameter [-] (for a full parameter table see Lawrence et al. (2018)), c_s is the CO_2 partial pressure at the leaf surface [Pa] and P_{atm} is the atmospheric pressure [Pa] and D is the vapor pressure deficit at the leaf surface [kPa].

The leaf transpiration is regulated by the leaf water potential. It is calculated for shaded and sunlit leaves separately based on maximum transpiration multiplied by the percent of maximum transpiration as modeled by the sigmoidal loss function (Equation 2) (Lawrence et al., 2018). Plant water stress is then calculated for shaded and sunlit leaves separately as the ratio of stomatal conductance of the leaf transpiration relative to maximum stomatal conductance corresponding to maximum transpiration (Lawrence et al., 2018). Leaf transpiration and the transpiration water stress (transpiration beta) are calculated for sunlit and shaded leaves separately as follows:

$$E = E_{max} \cdot 2^{-\left(\frac{\psi}{p50_e}\right)^{c_k}}, \tag{4}$$

$$\beta_t = \frac{g_s}{g_{s,max}}, \tag{5}$$

where E is leaf transpiration [mm/s], E_{max} is the leaf transpiration in the absence of water stress [mm/s], β_t is the transpiration water stress [-], g_s is the stomatal conductance of water corresponding to leaf transpiration [μmol/m^2/s], and $g_{s,max}$ is the stomatal conductance of water corresponding to maximum transpiration [μmol/m^2/s]. The calculated transpiration water stress is then used for the attenuation of photosynthesis, where $\beta_t = 1$ is no water stress and $\beta_t < 1$ is the relative transpiration water stress.

For a more detailed description of the algorithms applied in CLM5, the reader is referred to the model technical description (Lawrence et al., 2018, 2019) and references therein (the CLM5 technical model documentation can be accessed here: https://escomp.github.io/ctsm-docs/versions/release-clm5.0/html/tech_note/index.html).

In this study, all simulations were carried out with a modified version of CLM5 that was developed in an earlier study by Boas et al. (2021). This CLM5 version was extended with an adaptation of the winter cereal representation after Lu et al. (2017), a cover cropping and crop rotation routine, and crop phenology parameters for the northern hemisphere crop types winter wheat, sugar beet and potatoes. These modifications have proven to significantly improve the simulation of energy fluxes, vegetation states and carbon fluxes (e.g., leaf area index (LAI), net ecosystem exchange (NEE) and yield) at several Central European sites (Boas et al., 2021).

2.2. Study Areas and Input Data

Simulations were conducted for two intensive cropping regions in different climate zones, spanning a two-decade period from 1999 to 2019 (Figure 1). The first domain covers large parts of the south-east Australian wheat belt in the state of Victoria in Australia (AUS-VIC) and is characterized by large rain-fed agricultural areas with large paddock sizes, primarily dedicated to cereal cultivation, and extensive naturally vegetated or woody areas (i.e., grasslands, native woody cover and woody horticulture). The agricultural parts of the domain are mostly characterized by deep groundwater tables. The main cash crop in this domain is winter wheat, followed by barley and canola. Land cover information for the AUS-VIC domain was based on the Victorian Land Use Information

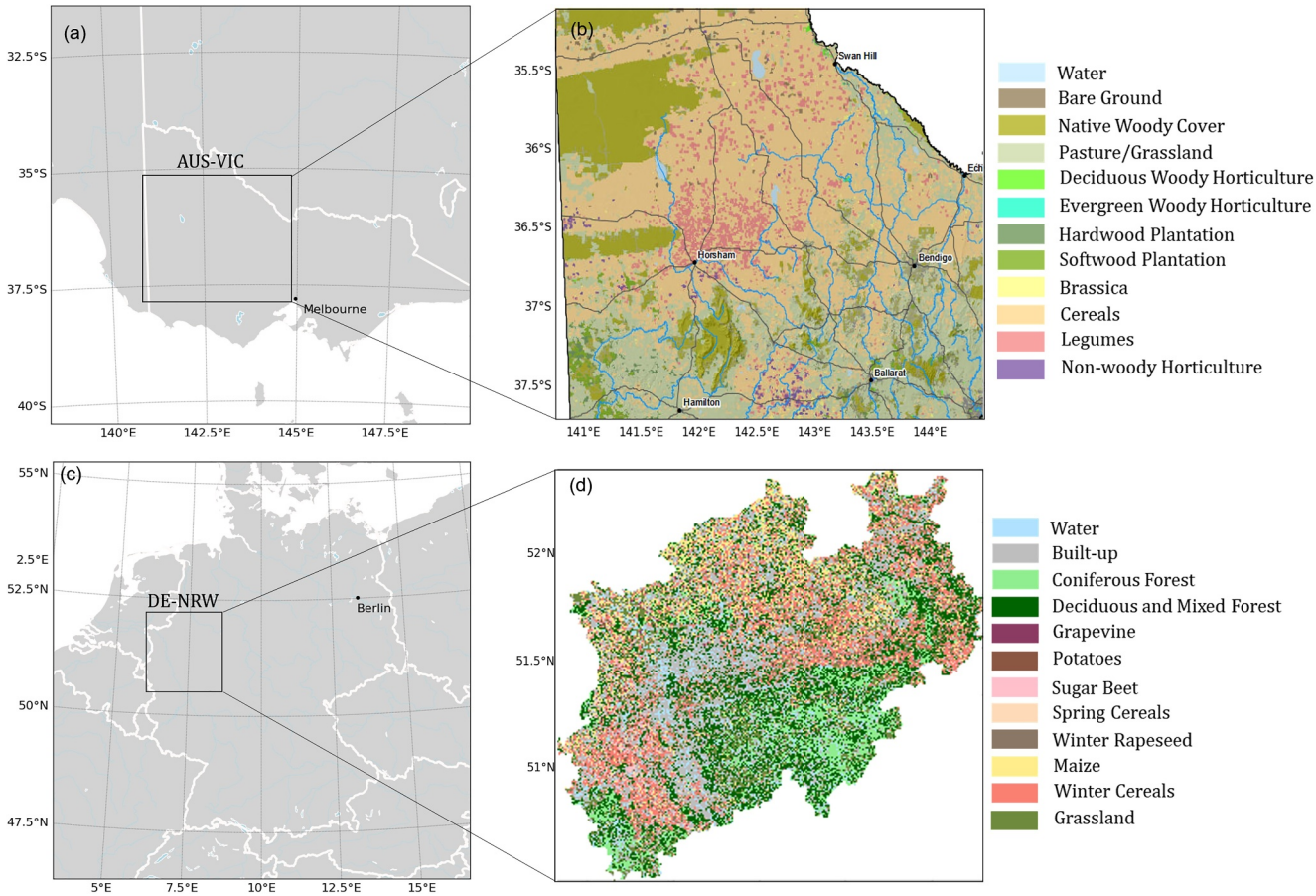

Figure 1. (a) AUS-VIC simulation domain extent and (b) dominant land use type based on VLUIS data, modified after (Victoria Government Data Directory, 2020); (c) DE-NRW simulation domain extent and (d) dominant land use type based on Griffiths et al. (2019), modified from Boas et al. (2023).

System (VLUIS) product for the year 2016 (Morse-McNabb et al., 2015; Victoria Government Data Directory, 2018). This data set was generated through a combination of time series analysis of remote sensing data (MOD13Q1 or MYD13Q1 by NASA) and annually collected field data (Morse-McNabb et al., 2015), and provides detailed information on land use and land cover for the entire state of Victoria. Unfavorable weather conditions for winter crop farming can have profound impacts on regional grain production and yield per area (ABARES, 2020).

The second domain covers the state of North Rhine-Westphalia in Germany (DE-NRW). It is characterized by a diverse landscape with urban, natural, and mixed agricultural areas that are mostly rain-fed. The groundwater regime in this domain is largely characterized by relatively shallow groundwater tables and a high degree of human influence due to intensive land use and urbanization. Land cover information for the DE-NRW domain was based on the 30-m resolution land cover data set by Griffiths et al. (2019). This data set was generated from Sentinel-2A MSI and Landsat-8 OLI observation data from the NASA Harmonized Landsat-Sentinel data set for the year 2016 (Claverie et al., 2018). Compared to agricultural reference data, the derived crop type and land cover map showed a high overall accuracy of over 80%, particularly for crop types with high abundances such as cereals, maize, and canola (Griffiths et al., 2019). The agricultural land cover in DE-NRW is primarily concentrated in the northern and western part of the domain, alongside natural vegetation and urban areas. The main cash crops in this region are winter wheat, winter barley, corn, sugar beet and rape seed (Figure 1, BMEL, 2020, 2022). In the southern part of the domain, which includes the Eifel, Bergisches Land and Sauerland regions, forests and grassland are the dominant land cover. Due to a late cold spell in late February/early March, agricultural yields in the area were significantly impacted in 2018. In addition, extreme heat and dry spells during both summers of 2018 and 2019 resulted in unusually high spatial variability of yield, particularly for cereals (BMEL, 2020, 2022; NRW State Government, 2020).

For both domains, soil variables such as clay, sand and organic matter content were derived from the global SoilGrids database (Hengl et al., 2017). SoilGrids provides soil information at seven depths (0, 0.05, 0.15, 0.30, 0.60, 1 and 2 m) at 250 m, 500 m and 1 km spatial resolution. Further soil properties, such as the saturated hydraulic conductivity and soil retention parameters are calculated within CLM5 based on the pedotransfer function after Cosby et al. (1984). Irrigation demand is dynamically calculated based on simulated soil moisture conditions. For irrigated croplands, the model assesses daily whether irrigation is needed. Irrigation is triggered when the crop leaf area index is greater than zero and available soil water drops below a specified threshold.

The amount of fertilizer is prescribed by crop functional types and varies spatially based on the LUMIP land use and land cover change time series (Lawrence et al., 2016). The LUMIP time series (1999–2019) provides land use and land cover data for prescribing fertilizer rates in the CLM5 model. Fertilizer application is adjusted spatially based on crop type and regional agricultural intensity, representing average historical rates. While spatial variability is accounted for, temporal changes in fertilization practices are not simulated beyond a repeated annual cycle, with nitrogen from industrial fertilizer applied consistently over 20 consecutive days each year. Manure nitrogen is uniformly applied at 0.002 kg N/m^2/year across all crop types.

2.3. Forcing and Validation Data

The simulations were forced with the bias-adjusted global reanalysis data set WFDE5 (Cucchi et al., 2020). The WFDE5 data set provides all meteorological variables that are needed to force CLM5 (i.e., precipitation, incoming shortwave and longwave radiation, atmospheric pressure, relative humidity, wind speed and temperature) at hourly time step for the period from 1979 to 2019, and at a 0.5° spatial resolution. The data set was generated using the WATCH Forcing Data (WFD) methodology (Cucchi et al., 2020) based on the ERA5 reanalysis product (Hersbach et al., 2020).

In order to validate and compare model results for crop yields, we used state-wide agricultural statistics. For Victoria, official records of annual crop yields are available for all simulated years from the database of the Australian Bureau of Agricultural and Resources Economics and Sciences (ABARES). The Australian crop report contains both realized and forecast growing area, production and yield for the major winter and summer crops on the Australian state level. The reports are produced with the participation of industry contacts and are released quarterly with the most recent estimates and updated forecasts.

For the DE-NRW domain, official agricultural records are available for 2005–2019 from IT.NRW (2022). The yield statistics are compiled for selected agricultural crops at the state and municipal levels in Germany. This information is gathered through an annual crop reporting process, where industry contacts with local expertise for the reporting areas, typically a municipality, provide harvest information and yield estimates for the year. Acknowledging the inherent uncertainties in these agricultural data sets, that is, stemming from variations in reporting methods, they provide valuable insights into the overall magnitudes and yearly trends of crop yield. As such, they serve as suitable validation data sets to assess the general quality of yield predictions in terms of magnitudes and inter-annual variations.

For validating simulated soil moisture contents in the top soil layers (up to 0.06 m depth), we used the CCI Soil Moisture-Combined data set, version 07.1 (ESA-CCI), from the European Space Agency's (ESA) Soil Moisture Essential Climate Variable (ECV) Climate Change Initiative (CCI) project (Dorigo et al., 2017; Gruber et al., 2017, 2019; Preimesberger et al., 2021). The ESA-CCI-SM combined product provides global daily volumetric soil moisture data at a spatial resolution of 0.25° from 1978 to 2021 (Dorigo et al., 2017; Gruber et al., 2017, 2019; Preimesberger et al., 2021). Additionally, we compared simulation results to the Soil Moisture Active Passive (SMAP) mission Enhanced Level-3 radiometer soil moisture product (SMAP L3) that is available since March 2015 and comprises soil moisture retrievals at a spatial resolution of 36 and 39 km (Entekhabi et al., 2016). When comparing daily simulation results with satellite-derived soil moisture retrievals, it is important to acknowledge the inherent uncertainties of these products, for example, due the impact of vegetation cover and surface characteristics, variations in satellite sensor characteristics, algorithm and model uncertainties etc., which may result in a distorted representation of daily soil moisture dynamics (Seo & Dirmeyer, 2022). Furthermore, it is crucial to recognize that these products represent near-surface soil moisture and can only serve as an indication for the conditions over the entire root zone.

Table 1
Overview of Conducted Simulation Experiments, Used Forcing's and Simulation Period, for Each Domain

	Description	Forcing	Simulation period
I	Spin-up	CRUNCEP	1901–2016, loop
II	Realistic land cover simulations	WFDE5	1999–2019
III	Winter wheat monoculture experiments	WFDE5	1999–2019
IV	Winter wheat monoculture experiments with reduced precipitation	modified WFDE5 with 50% reduced precipitation	1999–2019

2.4. Simulation Experiments and Analysis

As a first step, a model spin-up of more than 850 simulation years was conducted for each simulation domain so that the model ecosystem variables reach equilibrium prior to production simulations (experiment I in Table 1). This spin-up process comprised an initial phase under accelerated decomposition conditions lasting more than 300 years, followed by a final phase in normal mode lasting over 500 years. The spin-up is essential to attain equilibrium in ecosystem carbon and nitrogen pools, gross primary production, and total water storage. Notably, achieving equilibrium is particularly time-intensive for the slow carbon and nitrogen pools. For the spin-up, the combined global CRUNCEP atmospheric forcing data set (Viovy, 2018) was used, which consists of the CRU TS3.2 0.5 × 0.5° data set covering the period from 1901 to 2002 (Harris et al., 2014) and the NCEP reanalysis 2.5 × 2.5° 6-hourly data set available for 1948–2016 (Kalnay et al., 1996).

Then, simulations were conducted for both study domains with high resolution land cover information as described in Section 2.2 for the period 1999–2019, forced with WFDE5 data (experiment II in Table 1) (Cucchi et al., 2020). In the following step, the land cover was modified by setting the CFTs on all cropland land units within the domain to winter wheat. This synthetic winter wheat monoculture was simulated for both domains with the WFDE5 forcing data set, and for the years 1999–2019 (experiment III in Table 1). Additionally, the same set of simulations was conducted for the synthetic winter wheat monocultures with 50% reduced WFDE5 precipitation to synthetically create more drought stress for the crop (experiment IV in Table 1).

The model performance was statistically evaluated using the root mean square error (RMSE), the Pearson correlation coefficient (r), the squared correlation coefficient (R^2), and the mean bias error (MBE):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2}, \tag{6}$$

$$r = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{(n-1)\cdot\sigma_x\sigma_y}, \tag{7}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - x_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}, \tag{8}$$

$$\mathrm{MBE} = \frac{\sum_{i=1}^{n}(x_i - y_i)}{n}, \tag{9}$$

where n is the total number of time steps (days or years), x_i and y_i are the simulated and the observed values of a given variable at every time step i, overbar represents the mean value, and σ_x and σ_y are the standard deviations of the simulated and observed data respectively.

In order to quantify the inter-annual variability of annual crop yields and daily soil moisture contents, the mean absolute anomaly (MAA [%]), the mean absolute deviation (MAD) and mean absolute deviation ratio ($\mathrm{MAD_r}$ [-]) were defined as follows:

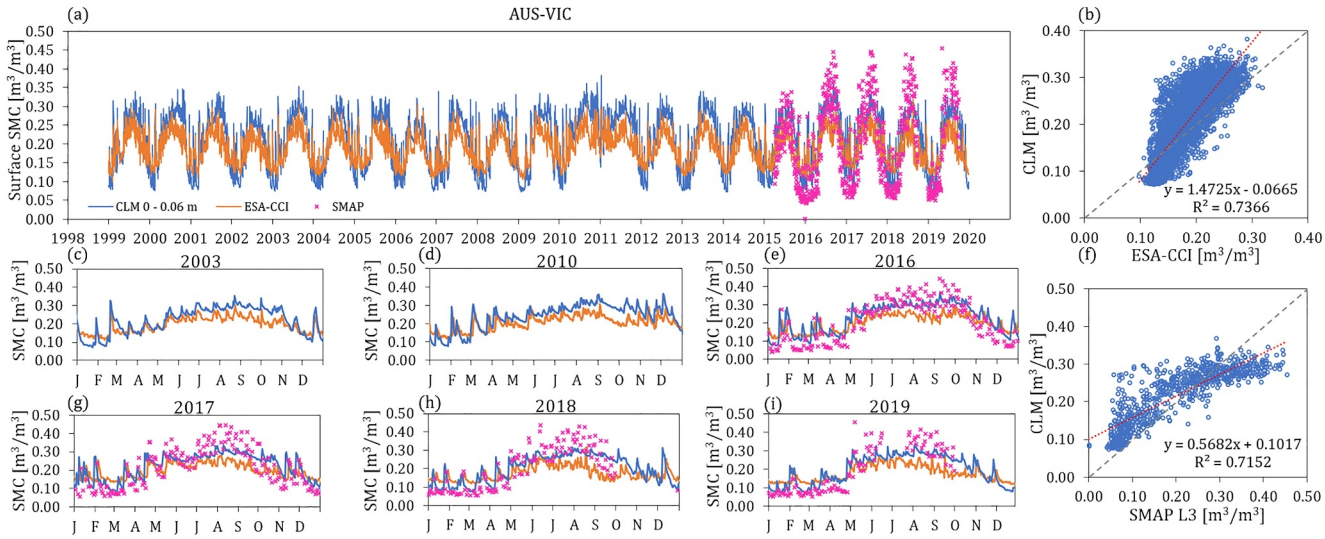

Figure 2. Simulated daily surface soil moisture (0–0.06 m depth) throughout the AUS-VIC domain (a) from 1999–2019, and (c, d, e, g, h, i) for individual years, compared to the ESA-CCI product and available SMAP L3 data. Scatterplots show the correlation between simulated SMC and (b) ESA-CCI, and (f) SMAP L3, with the respective regression equations.

$$\mathrm{MAA} = \frac{\sum_{i=1}^{n}\left(\left|\frac{Y_i - \overline{Y}}{\overline{Y}}\right| \cdot 100\right)}{n}, \tag{10}$$

$$\mathrm{MAD} = \frac{\sum_{i=1}^{n}\left(\left|Y_i - \overline{Y}\right|\right)}{n}, \tag{11}$$

$$\mathrm{MAD}_r = \frac{\mathrm{MAD}_X}{\mathrm{MAD}_y}, \tag{12}$$

where i is time step (days or years) and n the total number of time steps, Y_i is the respective variable [t/ha] at every time step i, $\overline{Y}$ is the 1999–2019 average of that variable [t/ha], $\mathrm{MAD_X}$ and $\mathrm{MAD_y}$ are the mean absolute deviation [t/ha] for the simulations and the reference data set respectively.

The absolute anomaly (AA) and the absolute deviation (AD) provide the percentage difference and the average absolute difference, respectively, between observations at the respective time step and the long-term average. Their mean values MAA and MAD provide insight into overall variability. Lower values in both metrics indicate less variability and a closer alignment to the long-term average. The $\mathrm{MAD_r}$ compares the calculated variability (MAD) between simulations and the reference data set, with values nearing one indicating greater similarity.

3. Results

3.1. Soil Moisture Regime

We compared the simulated surface soil moisture content (SMC) of the top soil layers (0–0.06 m depth) with data from the combined ESA-CCI product (Dorigo et al., 2017; Gruber et al., 2017, 2019) and the satellite derived SMAP L3 data (Entekhabi et al., 2016).

For AUS-VIC, the simulated SMC is systematically higher than ESA-CCI during the main cropping season (May–October) and lower during the austral summer months (December–February), resulting in a R^2 value of approximately 0.7 (Figure 2). The SMAP L3 product generally showed larger day-to-day fluctuations compared to the ESA-CCI data, with higher values during the main cropping season (May–October) and lower values during the summer months (December–February) than ESA-CCI. The simulations resulted in SMC very close to

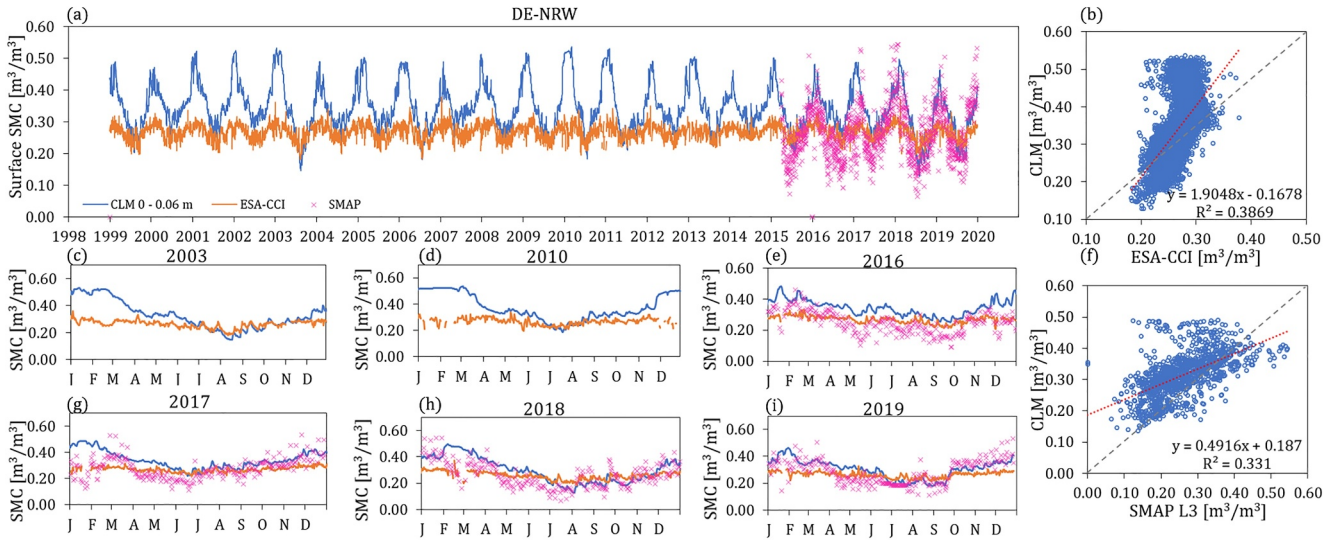

Figure 3. Simulated daily surface soil moisture (0–0.06 m depth) throughout the DE-NRW domain (a) from 1999–2019, and (c, d, e, g, h, i) for individual years, compared to the ESA-CCI product and available SMAP L3 data. Scatterplots show the correlation between simulated SMC and (b) ESA-CCI, and (f) SMAP L3, with the respective regression equations.

the SMAP L3 product during the first quarter of the year, while underestimating maximum values in the SMAP L3 data set during the growing season (May–October).

For the DE-NRW domain, comparison against ESA-CCI revealed large deviations compared to the simulated surface SMC, resulting in an R^2 value of approximately 0.4. For DE-NRW, the simulated winter season (December–February) and early growing season (March, April) SMC were significantly overestimated for all years, while late summer SMCs (August, September) were systematically underestimated with respect to ESA-CCI (Figure 3). The SMAP L3 product showed smaller values compared to ESA-CCI and was largely overestimated by CLM5 results for the majority of years and months, particularly during the main growing season (April–August) (Figure 3). The drought years of 2018 and 2019 are reflected in lower simulated SMC compared to the 1999–2019 average, in particular in (late) summer, where CLM5 showed lower values than SMAP L3 in 2019 (Figure 3i).

Simulated soil moisture by CLM5 showed a smaller inter-annual variability than SMAP, but a larger inter-annual variability than ESA-CCI. For the AUS-VIC domain, the mean absolute anomaly (MAA) for simulated soil moisture was approximately 30% (0.063 m³/m³ MAD), while it was 19% (0.036 m³/m³ MAD) for ESA-CCI and 48% (0.097 m³/m³ MAD) for SMAP (Table 2). Similarly, for the DE-NRW domain, the simulated soil moisture showed an MAA of approximately 17% (0.058 m³/m³ MAD), while ESA-CCI showed 7% MAA (0.02 m³/m³ MAD) and SMAP L3 had 23% MAA (0.058 m³/m³ MAD). The correlations between the simulated daily surface soil moisture and the reference data sets were generally higher for the AUS-VIC domain compared to DE-NRW. The correlation coefficient between simulated soil moisture and both reference data sets in AUS-VIC was approximately 0.85, whereas it was approximately 0.6 for both SMAP and ESA-CCI in DE-NRW.

3.2. Regional Crop Productivity

For the multi-decadal simulation runs we compared the simulated annual crop yields for DE-NRW and AUS-VIC to available yield records. We calculated yield variability as mean absolute deviation (MAD) and mean absolute anomaly (MAA) both for simulations and yield records, as well as the ratio of MAD from simulations and records (MAD_r) as explained in Section 2.4 (Table 3).

In general, the simulations produced mean annual crop yields that are comparable to the mean observed yields across both domains and for all considered crop types in terms of overall magnitudes (Figures 4a and 4c). For AUS-VIC, the average simulated mean annual crop yield (excluding sorghum) was 1.57 t/ha, compared to 1.75 t/ha in records. In DE-NRW, the simulated mean yield averaged for all considered crops was 6.7 t/ha, while the official records reported approximately 7 t/ha. Additionally, there were differences in the simulated yield amounts

Table 2
Mean Annual SMC, Mean Absolute Anomaly (MAA), Mean Absolute Deviation (MAD) and Mean Absolute Deviation Ratio (MAD_r) for Simulated Daily Soil Moisture Content, and Daily Soil Moisture Data From ESA-CCI and SMAP L3 From 1999–2019, for AUS-VIC and DE-NRW Respectively

	Mean SMC [m^3/m^3]	MAA [%]	MAD [m^3/m^3]	MADr [−]	r [−]	RMSE [m^3/m^3]	MBE [m^3/m^3]
AUS-VIC							
CLM	0.21	29.927	0.063	-	-	-	-
ESA-CCI	0.188	19.012	0.036	1.757	0.858	0.032	0.01
SMAP L3	0.203	48.038	0.097	0.645	0.846	0.014	0.001
DE-NRW							
CLM	0.344	16.869	0.058	-	-	-	-
ESA-CCI	0.267	7.041	0.019	3.087	0.622	0.064	0.033
SMAP L3	0.262	26.306	0.069	0.843	0.574	0.024	0.004

Note. Corresponding performance parameters r, RMSE and MBE were calculated with respect to the validation data sets ESA-CCI and SMAP L3.

and inter-annual yield variability between the two domains (Figures 4b and 4d). The simulated annual yield amounts were significantly higher in DE-NRW than in AUS-VIC for all crops, which is consistent with official yield statistics (Table 3). Additionally, the simulated inter-annual variability was lower for DE-NRW than for AUS-VIC, which is also the case in the official records (Figures 4b and 4d, Table 2).

For AUS-VIC, profound differences were observed in the inter-annual yield variability between the simulation results and crop survey records, with the absolute deviation (AD) being more than 3 times lower on average in simulation results compared to records (Figure 4b and Table 3). Specifically, the mean absolute anomaly was approximately 5% (0.12 t/ha MAD) for winter wheat in the simulations and 29% in the records. Similarly, the MAA and MAD values for sorghum were substantially underestimated, with 9% and 0.2 t/ha for simulations, compared to 59% and 1.3 t/ha for yield records. Overall, the lowest MAD ratios were reached for sorghum and winter wheat, with 0.15 and 0.2 respectively, while simulation results for canola and barley had ratios of 0.95 and

Table 3
Mean Annual Crop Yield, Mean Absolute Anomaly (MAA), Mean Absolute Deviation (MAD) and Mean Absolute Deviation Ratio (MAD_r) for Simulated and Recorded Yields for 1999–2019, Averaged for Winter Wheat, Barley, Canola and Sorghum for AUS-VIC, and for Winter Wheat, Spring Wheat, Canola and Corn for DE-NRW

	Mean yield [t/ha]		MAA [%]		MAD [t/ha]					
	Obs	CLM	Obs	CLM	Obs	CLM	MAD_r [−]	r [−]	RMSE [t/ha]	MBE [t/ha]
AUS-VIC										
Winter wheat	1.95	2.25	28.65	5.26	0.56	0.12	0.21	0.39	0.68	0.3
Barley	1.97	1.31	28.75	22.62	0.57	0.29	0.51	0.07	0.98	−0.64
Canola	1.32	1.15	21.99	24.03	0.29	0.28	0.95	0.11	0.52	−0.18
Sorghum	1.84	2.2	59.21	9.05	1.34	0.2	0.15	0.07	1.76	0.42
Mean[a]	1.75	1.57	34.65	15.24	0.69	0.22	0.32	0.12	-	-
DE-NRW										
Winter wheat	8.32	7.9	5.39	2.63	0.45	0.27	0.6	0.42	0.61	−0.4
Spring wheat	6.3	5.09	8.29	7.27	0.52	0.43	0.83	0.11	1.4	−1.12
Canola	3.87	5.07	5.05	8.49	0.2	0.49	2.51	0.22	1.39	1.28
Corn	9.97	8.76	7.93	3.43	0.79	0.3	0.38	0.13	1.58	−1.29
Mean	7.11	6.7	6.67	5.46	0.49	0.37	0.76	0.12	-	-

Note. Corresponding performance parameters r, RMSE and MBE were calculated for the annual mean yield from 1999 to 2019 for AUS-VIC and 2005–2019 for DE-NRW with respect to available observations. [a]Sorghum excluded.

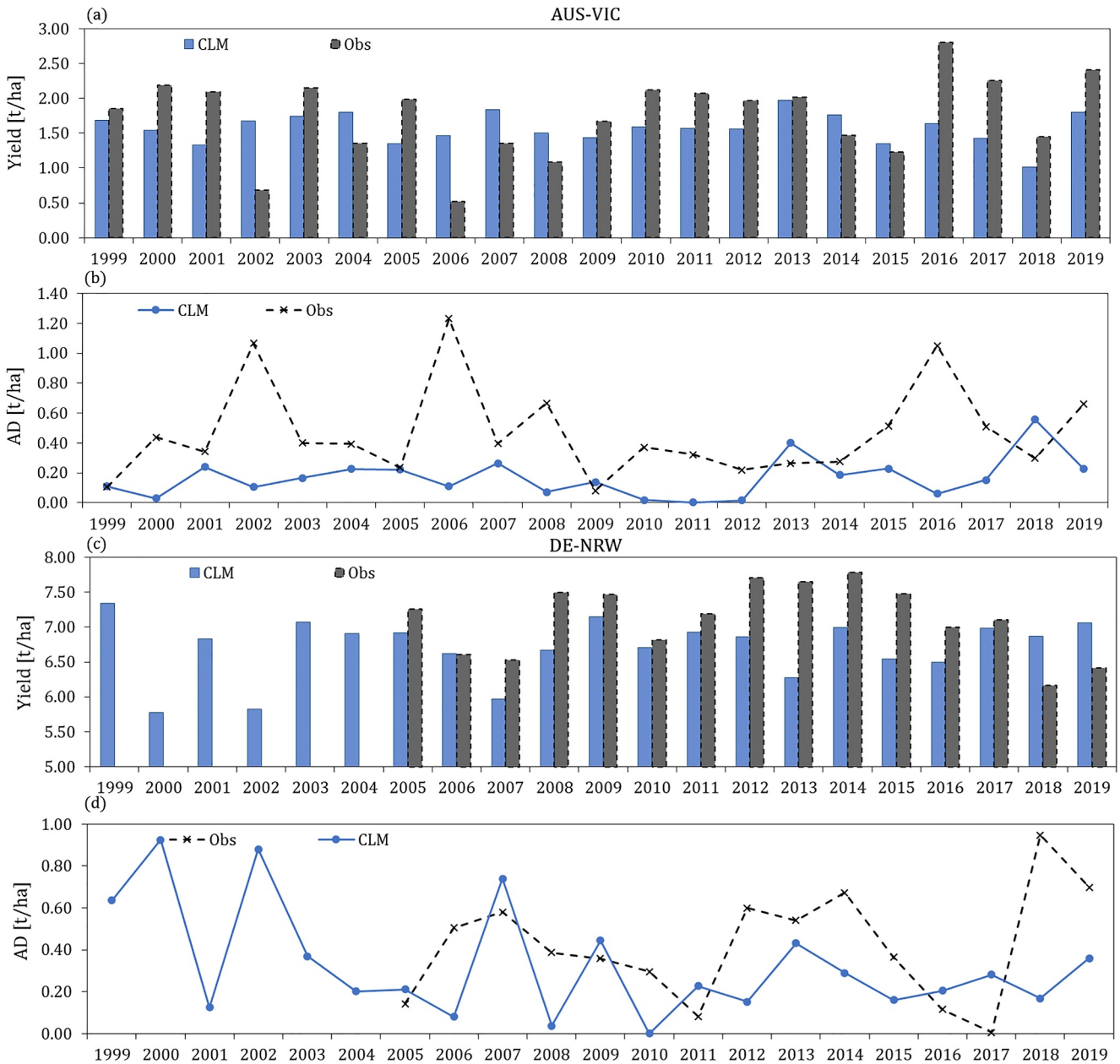

Figure 4. Simulated mean annual crop yield (CLM) and recorded mean annual yield (Obs), and corresponding absolute deviation from the 1999–2019 mean for each year (AD), averaged for all regarded crops, for (a–b) the AUS-VIC domain (winter wheat, barley, canola, sorghum) and (c–d) the DE-NRW domain (winter wheat, spring wheat, canola, corn).

0.5 respectively (Table 3). The crop functional types (CFTs) for barley and canola are both derived from the spring wheat CFT (with adjusted values for several phenological parameters such as maximum LAI, maximum crop height, etc.), which explains the very similar yield predictions and inter-annual fluctuations for these two crops (Figures 5c and 5e). The drought year of 2018 and the reduction in recorded yields compare to previous years was only captured in the simulations for barley and canola (Figures 5c and 5e).

For the DE-NRW domain, the simulated total annual crop yield was close to official records, especially for winter wheat (Figure 6a, Table 3). However, simulated inter-annual yield variability was considerably lower than in the yield records for winter wheat (with a MAA of 2.6% and MAD of 0.3 t/ha in the simulations compared to 5.4% and 0.5 t/ha in the records) and corn (MAA of 3.4% and MAD of 0.3 t/ha in the simulations, and 7.9% and 0.8 t/ha in the records, approximately), resulting in low MAD ratios of 0.6 and 0.4, respectively (see Figure 6 and Table 3).

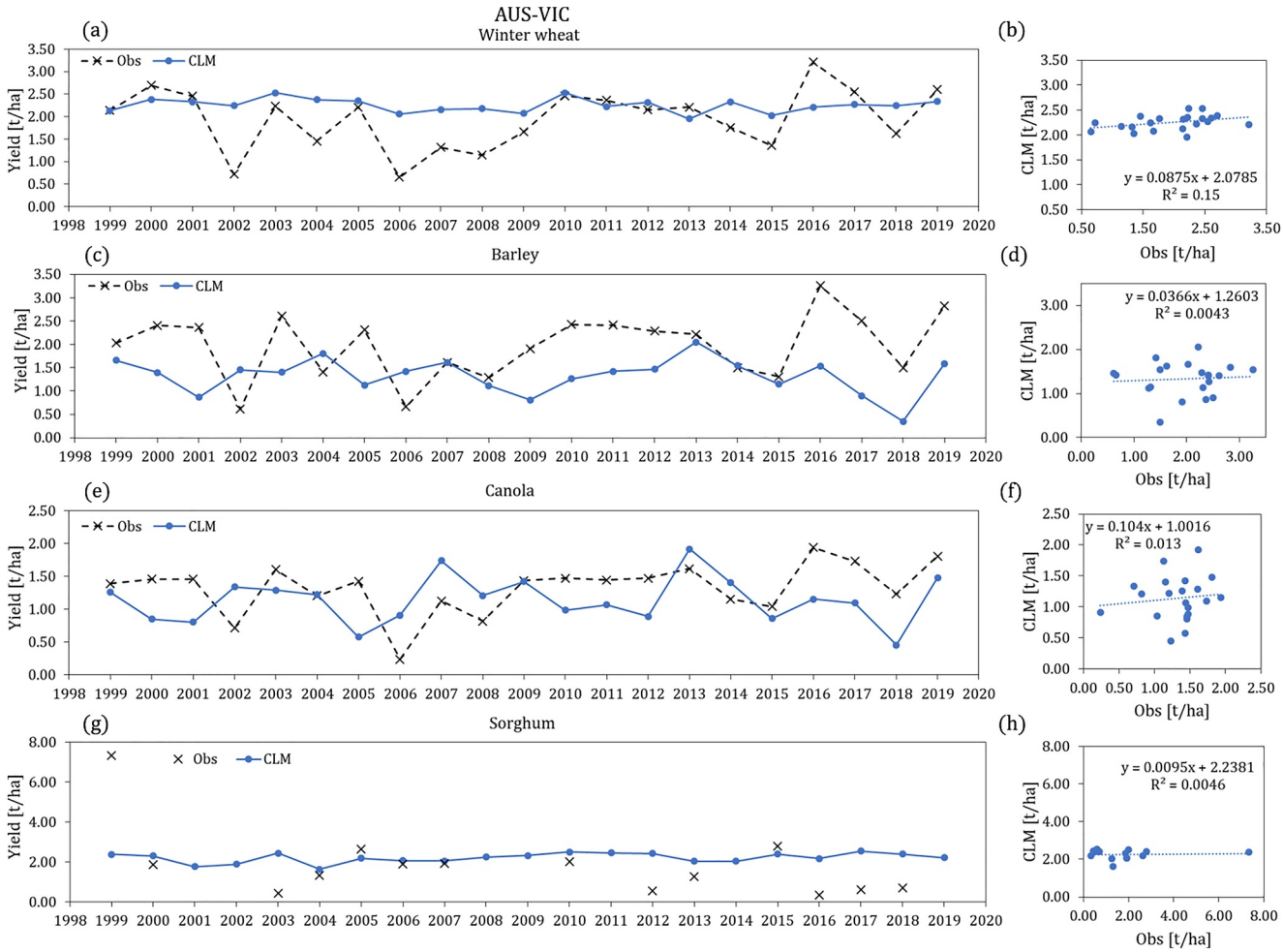

Figure 5. (left) Simulated mean annual crop yields for (a) winter wheat, (c) barley, (e) canola and (g) sorghum from 1999–2019 throughout the AUS-VIC domain compared to available records from ABARES (2022) with (right) corresponding correlations between simulated and observed values, with the respective R^2 values and regression equations. The corresponding data is also shown in Table A1.

Similar to the results for AUS-VIC, the influence of the shared crop parameterization between canola and spring wheat was evident in the simulation results for D-NRW. While the inter-annual yield variations for spring wheat were reasonably well-captured in the simulations, with a MAD ratio of approximately 0.8, the inter-annual yield variations for canola were overestimated in comparison to records, resulting in a very high MAD ratio of 2.5.

Throughout AUS-VIC, crop yield is strongly correlated with the total amount of rainfall throughout the cropping season (May–October), as demonstrated by the positive correlation of total growing season rainfall and recorded grain yields (Figure 7). This relationship was not evident in the CLM simulation results. In addition, a weak correlation could be observed between the simulated annual yield and the simulated root zone soil moisture (0.02–0.32 m depth) (Figure 8a).

For DE-NRW, the correlation between total growing season rainfall and recorded or simulated crop yields is not significant, reflecting the energy-limited regime in the area. This is demonstrated by the weak correlation between recorded yields and seasonal rainfall amounts. For spring wheat, the influence of precipitation is reflected in the correlation with recorded yields (wetter years resulted in higher grain yield), while for other crops the constraining effects of water availability and energy are more in balance. The simulation results showed a slightly negative correlation of yield, particularly pronounced for spring wheat and canola, with seasonal rainfall (and root zone soil moisture), which might be related to less energy input (smaller global radiation) for growing seasons with higher precipitation amounts (Figure 8b). With this negative correlation the simulation results overestimated the effects of energy limitation for the domain.

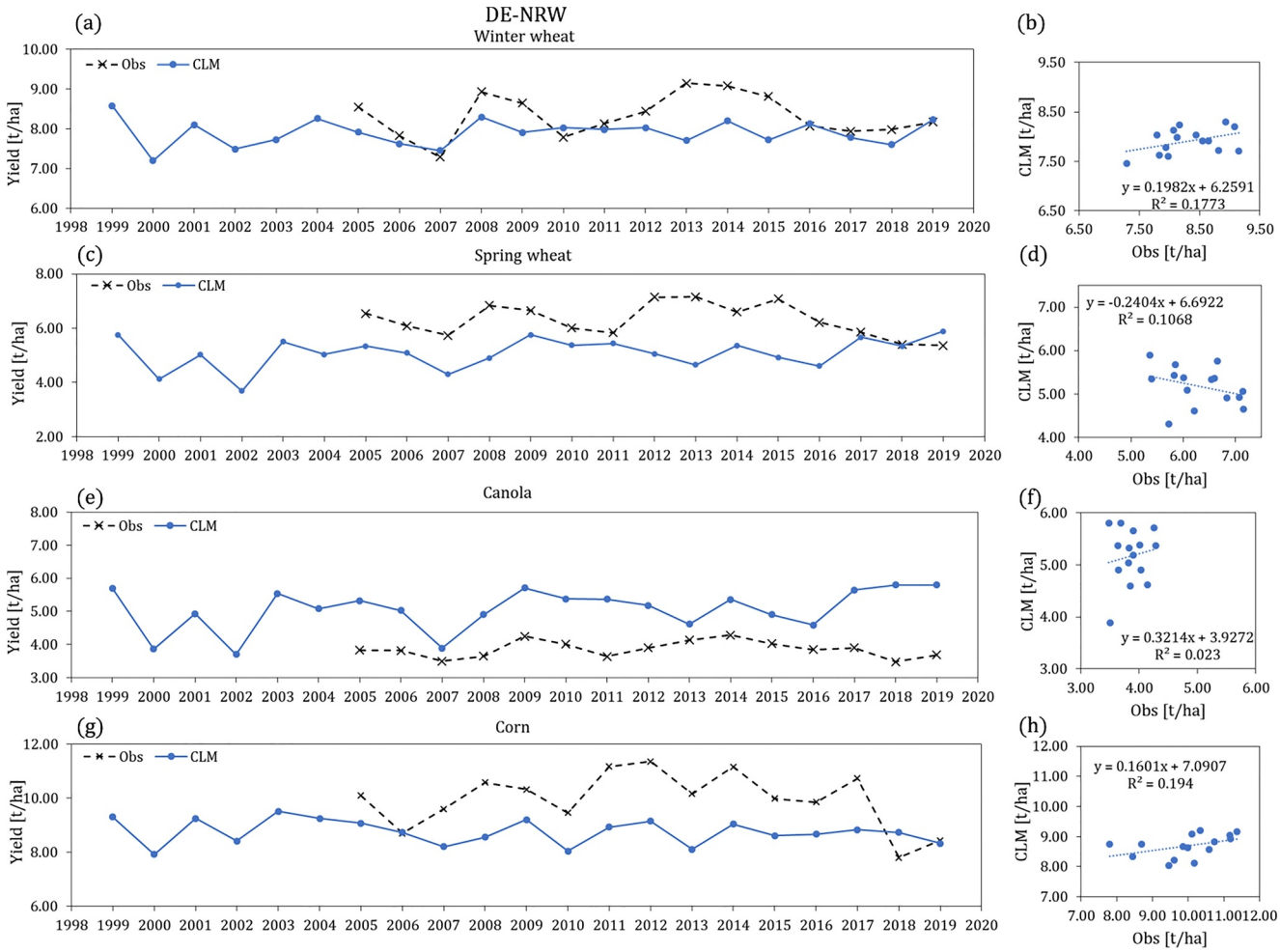

Figure 6. (left) Simulated mean annual crop yields for (a) winter wheat, (c) spring wheat, (e) canola and (g) corn from 1999–2019 in the entire DE-NRW domain compared to available records from 2005–2019 (IT.NRW, 2022) and (right) corresponding correlations between simulated and observed values, with the respective R^2 values and regression equations. The corresponding data is also shown in Table A2.

In order to better isolate the impact of soil moisture and precipitation on simulated crop yield, we additionally performed a multiple regression analysis, taking global (shortwave) radiation into account. We examined two different models to explain the simulated mean annual crop yield using Equation 1 the simulated mean seasonal root zone soil moisture and the mean seasonal global radiation, and Equation 2 the seasonal precipitation amount and the mean seasonal global radiation as independent variables. For DE-NRW, the model showed a moderate relationship between the variation in crop yield and changes in root zone soil moisture and global radiation with a R^2 of 0.38, and a slightly lower correlation of seasonal precipitation amount and global radiation with crop yield with a R^2 of 0.35 (Tables A3 and A4). For DE-NRW, global radiation exhibited a positive relationship, implying that increased radiation was associated with higher crop yields, while both the root zone soil moisture and seasonal rainfall amount exhibited negative relationships (not showing any statistical relevance with p-values >0.05) (Table A3). For the AUS-VIC domain, the model showed a weaker relationship between variations in crop yield and changes in root zone soil moisture and global radiation than in DE-NRW, with an R^2 of 0.16. Both the root zone soil moisture (p-value of 0.09) and global radiation (p-value of 0.26) were not statistically significant predictors in this case. The correlation of seasonal precipitation amount and global radiation with crop yield was slightly higher with an R^2 of 0.34, with both showing a positive relationship with crop yield (Tables A3 and A4).

Analysis of the simulated transpiration beta factor, which represents plant water stress in CLM5, did not show any apparent correlation with the simulated yield in both domains (Figure 9). The transpiration water stress ($\beta t = 1$ indicates absence of water stress, declining values indicate growing water stress) is utilized in CLM5 to regulate

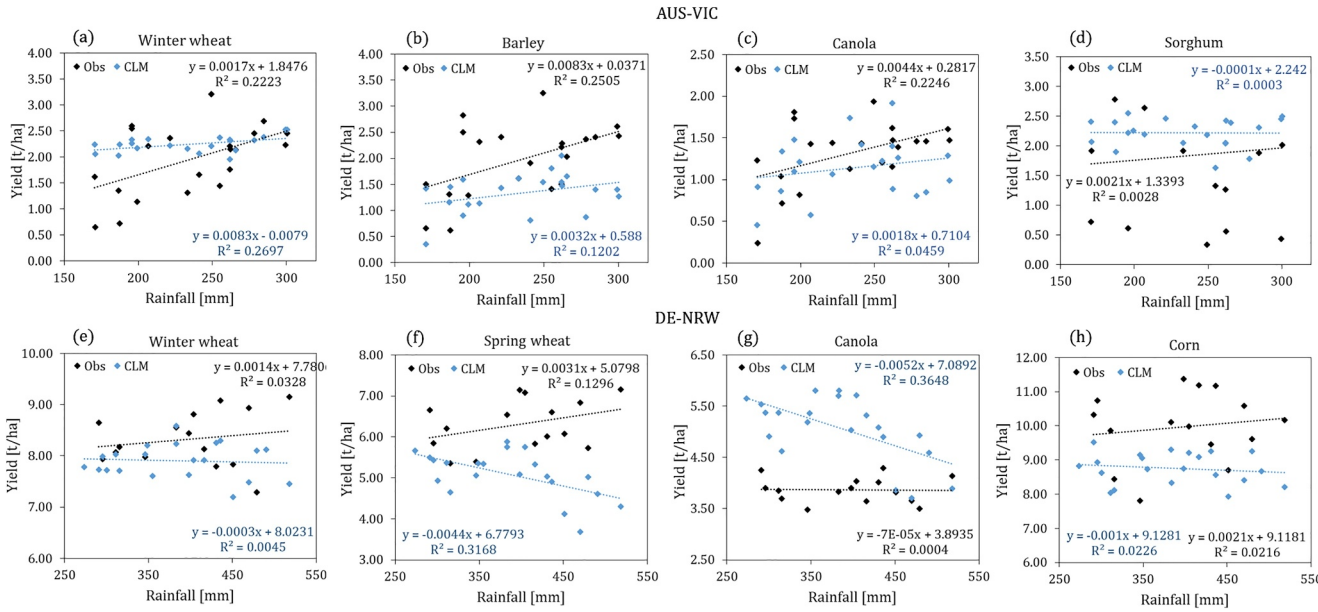

Figure 7. Relationship between mean annual crop yield (simulated (CLM) and recorded (Obs)) and WFDE5 cropping season rainfall amounts for (a–d) AUS-VIC (May–October) and (e–h) DE-NRW (April–September) spanning 1999–2019. The yields of the following individual crops are given: (a) winter wheat; (b) barley; (c) canola; (d) sorghum; (e) winter wheat; (f) spring wheat; (g) canola; (h) corn. The respective R^2 values and regression equations are given for (blue) simulation results and (black) records. Corresponding statistical results for the relationship between mean annual crop yield and seasonal rainfall are listed in Table A3.

plant photosynthesis. The lack of a correlation for AUS-VIC suggests that the simulated crop yield was not influenced by transpiration water stress (Figure 9a). In DE-NRW, even a slightly negative correlation could be observed for simulated crop yield and transpiration beta (Figure 9b). This is consistent with the results regarding the correlation of simulated yield with both root zone soil moisture and precipitation as discussed above (Figures 6 and 7).

3.3. Winter Wheat Monoculture Experiments With Reduced Precipitation

In a subsequent step, we conducted synthetic monoculture experiments where winter wheat cultivation was implemented exclusively across all crop land units. These experiments were performed using two different forcings: the default WFDE5 data set (CLM_WFDE5) and the WFDE5 data set with 50% reduction in

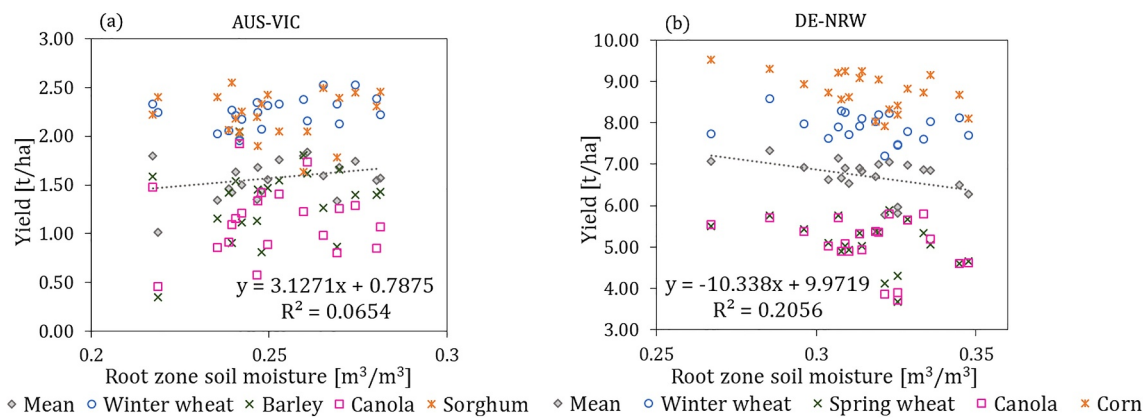

Figure 8. Relationship between mean annual crop yield (CLM) and the mean simulated root zone soil moisture (0.02–0.32 m depth) for the cropping seasons from 1999 to 2019, averaged for (a) AUS-VIC (May–October) and (b) DE-NRW (April–September), and for the respective crops. The respective R^2 values and regression equations for the mean yield are indicated in black. Corresponding statistical results are listed in Table A3.

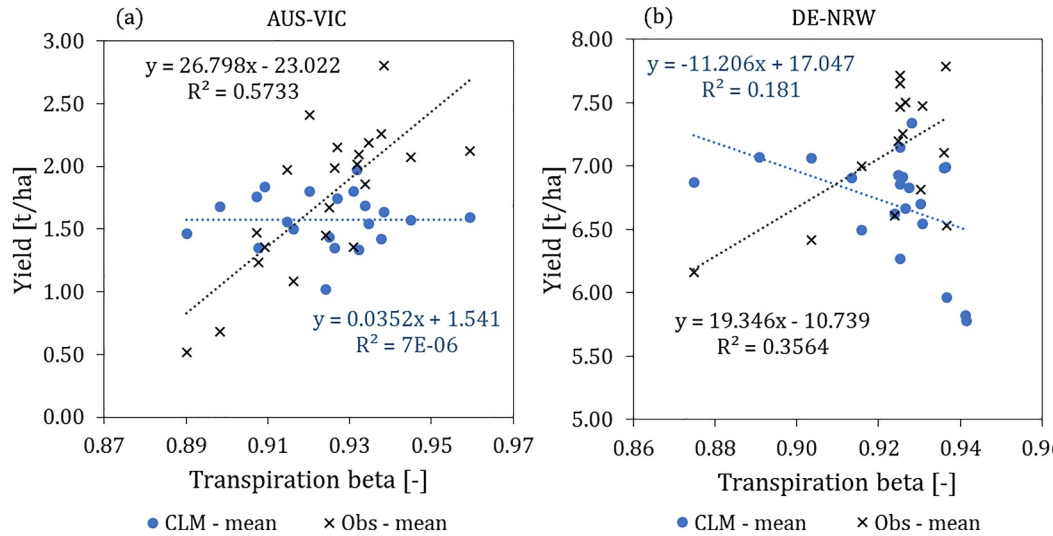

Figure 9. Relationship between simulated mean annual transpiration beta and simulated mean annual crop yield (CLM—mean), and recorded mean annual yield (Obs—mean) respectively, averaged for all regarded crops, for (a) the AUS-VIC domain (winter wheat, barley, canola, sorghum) and (b) the DE-NRW domain (winter wheat, spring wheat, canola, corn). The respective R^2 values and regression equations for (blue) simulations and (black) records are also given. Corresponding statistical results are listed in Table A3.

precipitation (CLM_LowP). The reduction in rainfall synthetically increased the plant water stress for both domains, as represented by the transpiration beta factor.

In the AUS-VIC domain, the simulations with reduced precipitation not only led to decreased mean yield amounts, introducing a larger negative bias in overall yield results, but also increased inter-annual variability compared to the WFDE5 simulations with unchanged precipitation. The recorded winter wheat yields exhibited a mean absolute anomaly of 28.7% (0.56 t/ha MAD), which was underestimated in the default CLM_WFDE5 runs with a mean absolute anomaly of 8.4% (0.16 t/ha MAD). In the reduced precipitation runs this underestimation was less pronounced with a mean absolute anomaly of 13.5% (0.18 t/ha MAD) (Figure 10b, Table 4). In the rain-fed and water-limited regions of the AUS-VIC domain, we expect a positive correlation between seasonal rainfall amounts and crop productivity. CLM results were able to capture this relationship only with the reduced precipitation, indicating an underestimation of plant water stress in the scenario with the default WFDE5 forcing (Figure 11a).

In the DE-NRW domain, the reduced precipitation runs consistently underestimated the total annual winter wheat yield for all years, except 2007 (Figure 10c and Table 4). The variability of yield increased from 3.3% (0.25 t/ha MAD) in default WFDE5 simulations to 7.5% (0.55 t/ha MAD) with reduced precipitation amounts (CLM_LowP), which was even higher than the recorded yield variability with 5.39% MAA (0.45 t/ha MAD). The overestimation of yield variability in the CLM_LowP runs was also reflected in the high MAD_r value of 1.2. By reducing precipitation amounts, we artificially created a more water-limited regime in the DE-NRW domain, resulting in a positive linear correlation between seasonal rainfall amounts and yield, which was even more pronounced than for the official yield records (Figure 11d).

The synthetic winter wheat monoculture experiments additionally offer the possibility to study simulated spatial differences as well as the effects of soil properties on simulated crop productivity in more detail (Figure 12). The highest variability of simulated annual crop yield is reached for regions with high sand contents, while in regions with relatively high clay contents, the inter-annual variability of yield is comparably low. This observed pattern suggests a link between the simulated yield and the higher water-retaining capabilities of clay-rich soils. In summary, the reduced precipitation amounts illustrate a more realistic water stress response from the crops, leading to a more realistic correlation of yield and rainfall amounts for the water-limited regime in AUS-VIC and higher inter-annual yield variabilities. However, the reduced precipitation lead to an underestimation of annual yield amounts for both domains compared to records (Tables A1 and A2).

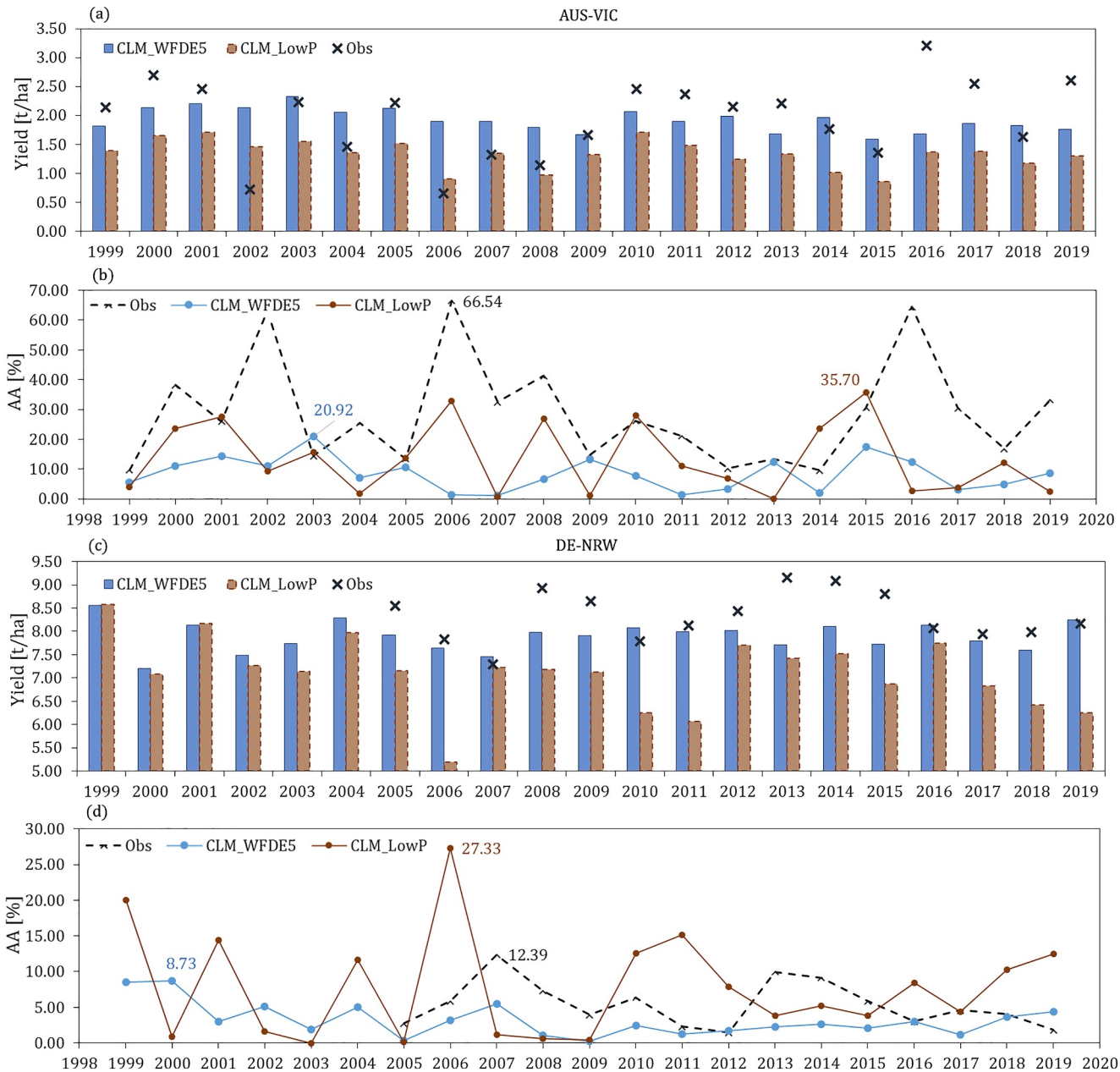

Figure 10. Mean annual crop yield from simulations forced with WFDE5 (CLM_WFDE5) and 50% reduced WFDE5 precipitation (CLM_LowP), and recorded mean annual yield (Obs) for (a) the AUS-VIC domain and (c) the DE-NRW domain, with (b, d) the corresponding annual absolute yield anomaly (AA) for each simulation scenario and domain. The corresponding data is also shown in Tables A1 and A2.

4. Discussion

During the growing season, the vast majority of agricultural land is under water-limited conditions (Koster et al., 2009; Nemani et al., 2003; Papagiannopoulou et al., 2017). Studies indicate a widespread trend of ecosystems moving from energy to water limited due to climate change (Denissen et al., 2022; Orth et al., 2023). Unsustainable water use in large parts of the world additionally increases the ecosystem vulnerability to drought with depleting groundwater and surface water resources (Samaniego et al., 2018; Taylor et al., 2013; Wada et al., 2010, 2012). Thus a reliable representation of the plant water stress regime and drought responses of vegetation are essential for the relevance of LSM applications for climate change research. The representation of

Table 4

Mean Annual Crop Yield, Mean Absolute Anomaly (MAA), Mean Absolute Deviation (MAD) and Mean Absolute Deviation Ratio (MAD_r) for Available Records and Results From Winter Wheat Monoculture Experiments, Forced With WFDE5 Precipitation (CLM_WFDE5) and 50% Reduced WFDE5 Precipitation (CLM_LowP), Averaged for 1999–2019, for Both Domains

	Mean yield [t/ha]	MAA [%]	MAD [t/ha]	MADr [−]	*r* [−]	RMSE [t/ha]	MBE [t/ha]
AUS-VIC							
Obs	1.95	28.65	0.56				
CLM_WFDE5	1.92	8.41	0.16	0.29	0.05	0.67	−0.03
CLM_LowP	1.33	13.51	0.18	0.32	0.59	0.83	−0.62
DE-NRW							
Obs	8.32	5.39	0.45				
CLM_WFDE5	7.89	3.20	0.25	0.56	0.29	0.64	−0.43
CLM_LowP	7.14	7.48	0.55	1.22	0.45	1.60	−1.46

Note. Corresponding performance parameters *r*, RMSE and MBE were calculated for the annual mean yield from 1999 to 2019 for AUS-VIC and 2005–2019 for DE-NRW with respect to available observations.

vegetation responses to changes in available water is a major issue for studying the impact of climate change on ecosystems and food security.

Our validation of simulated crop yields against state-wide statistics revealed reasonable estimates of total crop yield but underestimated inter-annual variability. Notably, better representation of inter-annual yield variability

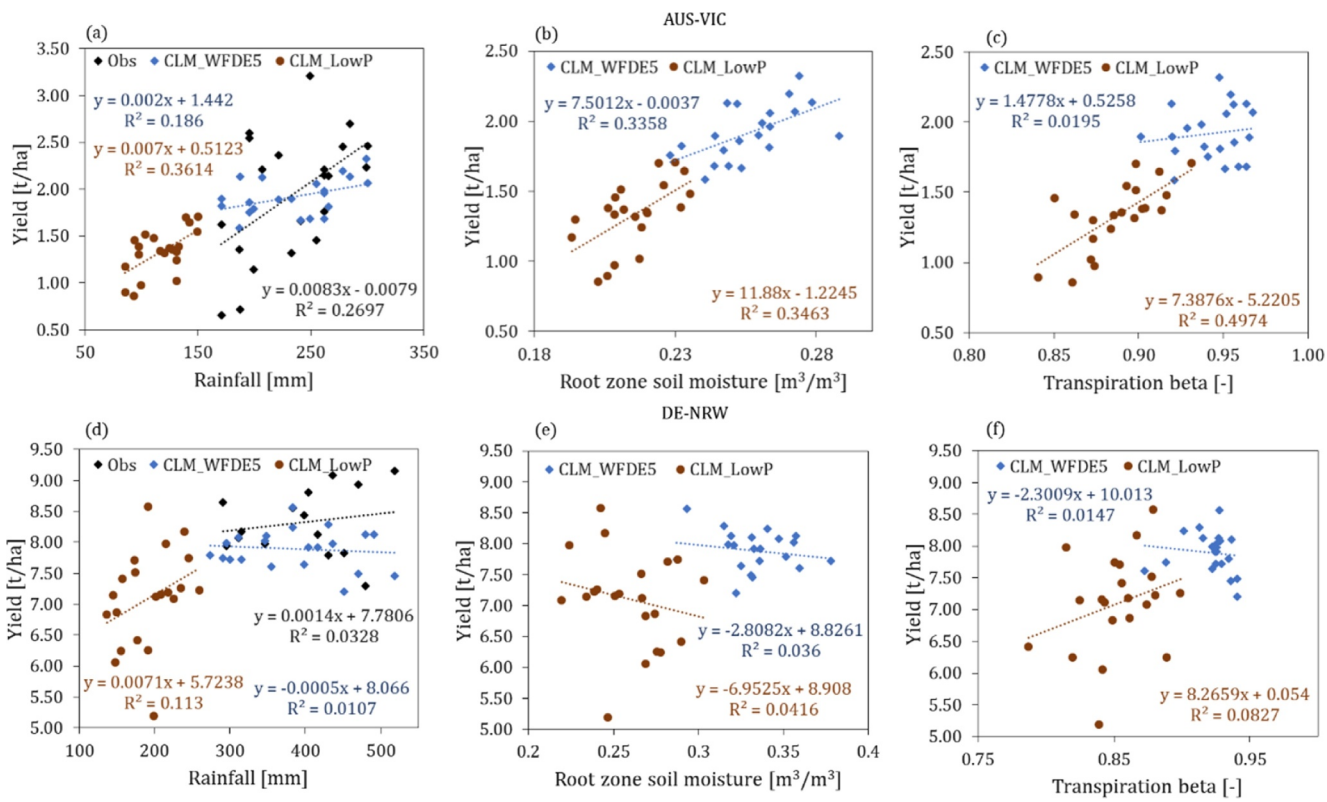

Figure 11. Relationship between mean annual winter wheat yield (either WFDE5 precipitation simulations (CLM_WFDE5), 50% reduced WFDE5 precipitation simulations (CLM_LowP) or records (Obs)) and the corresponding rainfall amounts for the cropping seasons of 1999–2019 (a, d), the simulated root zone soil moisture (b, e) and the simulated transpiration beta factor (c, f). Results are provided for the AUS-VIC domain (May–October) (a–c) and the DE-NRW domain (April–September) (d–f). The corresponding regression equations are indicated and color coded. The recorded yield is compared to WFDE5 annual precipitation.

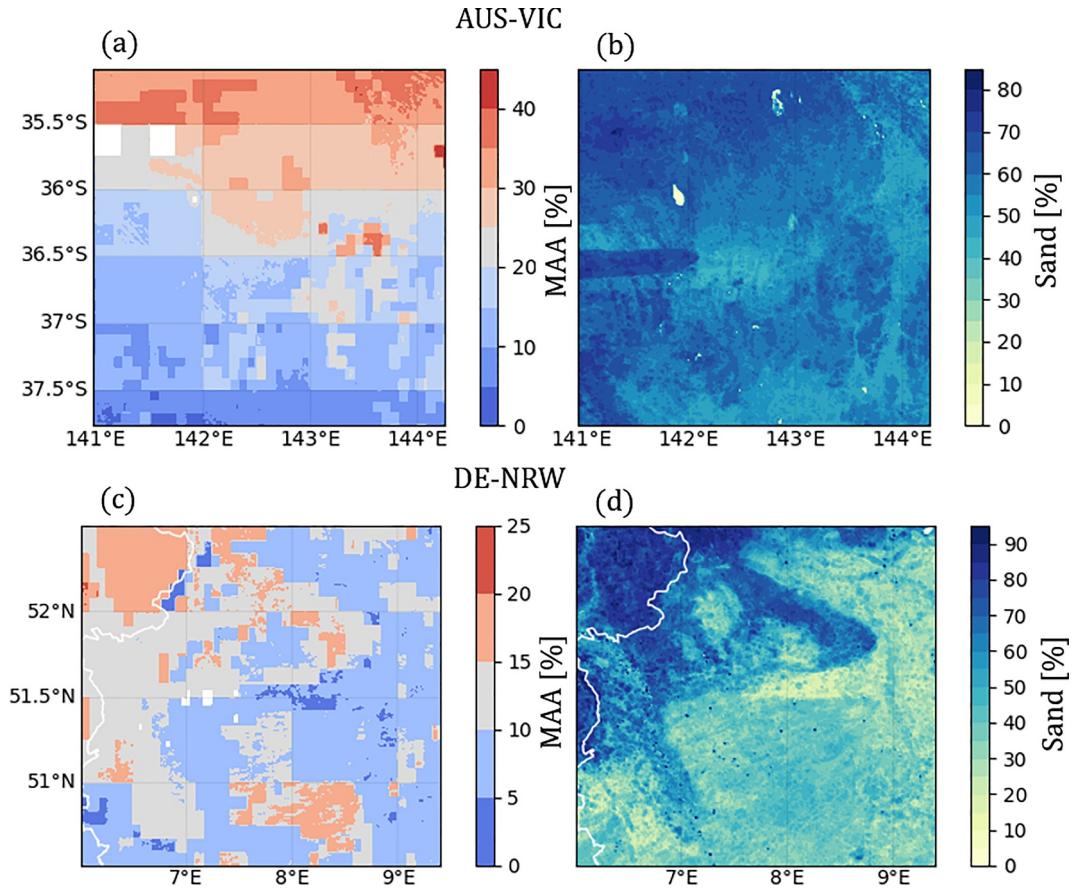

Figure 12. Spatial mean absolute yield anomaly (MAA) for winter wheat monoculture simulations (1999–2019) throughout (a) AUS-VIC and (c) DE-NRW, and (b, d) the sand content in the root zone throughout the respective domains, based on SoilGrids (Hengl et al., 2017).

was observed for CFTs related to spring wheat, which has been widely tested at global scale and its parameterization calibrated for CLM5 (Lombardozzi et al., 2020), including barley and canola, which exhibited similar magnitudes and fluctuations in annual yield. In addition, the model performed relatively well for winter wheat in the DE-NRW enhancements validated for winter wheat at several European sites in previous studies (Boas et a., 2021). In general, CLM5 is tuned to replicate average crop yields through parameter adjustments. Theoretically, well-calibrated parameters have the potential to offset the impacts of environmental factors such as pests, diseases, wildlife damage or pollutants. Previous studies highlighted the importance of including crop-specific physiological parameters in improving the quantification of the diurnal energy partitioning and yield estimates (e.g., Boas et al., 2021; Lu et al., 2017; Sulis et al., 2015). Further improving the parameterization within the crop phenology module of CLM5 for the individual crops could therefore help to alleviate model limitations.

Analysis of the correlation between simulated annual yield and recorded rainfall amounts, as well as simulated root zone soil moisture contents, revealed model limitations in representing the water-limited regime in AUS-VIC. Throughout Victoria, crop growth and yield is highly influenced by rainfall patterns and amounts (French & Schultz, 1984a, 1984b) which is also reflected in a strong positive correlation between total growing season rainfall amounts from reanalysis and observed total yields (Figures 7a–7d). This relationship is not reflected in CLM5 simulations. Only in reduced precipitation experiments for winter wheat, did we observe a positive linear relationship between yield and rainfall amounts; and yield and soil moisture for AUS-VIC (Figures 11a–11c). In the DE-NRW domain, the shallower groundwater tables and better water storing properties of soils lead to an absence of positive correlation between rainfall and overall annual yield compared to AUS-VIC. In addition, due to the high annual precipitation in general, soil water does not represent the main limiting factor for state-wide agricultural yields throughout DE-NRW, and global radiation is the most limiting factor.

Even during the drought year of 2018, the state-wide yield statistics for DE-NRW do not show a strong declining trend (Figure 4c). While severe yield losses were recorded in specific regions within the state NRW, other regions with shallower groundwater tables and soils with higher water-storing capacities experienced record yields due to more sunshine hours (which is also represented in the negative correlation between seasonal rainfall and yield, Figures 7e–7h). The slightly negative correlation between simulated crop yield and seasonal rainfall, as well as root zone soil moisture, in DE-NRW (which can be attributed to the reduction in sunshine hours caused by cloud cover associated with increased rainfall) suggests an exaggeration of the energy-limited regime by the land surface model and thus, a failure to correctly reproduce the effects of soil water stress on crop yield. The simulation results do not reflect the differences in the water-limited and energy-limited regimes between the two domains, except for a weak positive correlation between yield and rainfall in AUS-VIC, which indicates an inaccurate representation of the water-limited regime in AUS-VIC and a systematic model insensitivity toward drought stress. This is also underlined by the results from the multiple regression analysis (i.e., root zone soil moisture, precipitation and global radiation) that did not identify seasonal soil moisture as a relevant predictor for simulated crop yield in the AUS-VIC domain (Table A3).

Our results are consistent with other studies that indicate systematic issues across LSMs in simulating vegetation responses under drought conditions (e.g., De Kauwe, Zhou, et al., 2015; Trugman et al., 2018; Ukkola et al., 2016). Especially in water-limited agricultural regions, the selection of crop varieties is largely influenced by their resilience to prolonged dry periods. While plant physiological properties play an important role in the energy partitioning at the land surface and carbon fixation and drought sensitivity varies considerably among plants, state-of-the-art LSMs currently assume the same drought sensitivity for all crop types. The studies by Trugman et al. (2018) and Sulis et al. (2019) emphasize the importance of including mechanistic water limitation algorithms in LSMs to improve the representation of plant hydraulics and thus projections of the land carbon sink. Trugman et al. (2018) found that soil moisture-limited productivity and its uncertainties significantly influenced carbon cycle simulations and thus concluded that the representation of soil moisture represents a major source of uncertainty in land surface models.

In CLM5, the empirical soil moisture stress formulation from earlier model versions was replaced with a plant hydraulic stress model (Lawrence et al., 2019). However, one of the main challenges for the application of plant hydraulic models for different biomes is the parameterization. In CLM5, plant hydraulics are physically constrained by plant-dependent parameters, such as the conductivities of the soil-root interface and at the interfaces between each of the plant elements (Kennedy et al., 2019). Due to a lack of readily available data for these parameters, they do not vary for the different crops and represent first estimates, which is a challenge the model developers are well aware of Kennedy et al. (2019) and Lawrence et al. (2019). The simulated plant water states in CLM5 are physical properties that can be validated against field observations (e.g., Konings et al., 2017; Li et al., 2017) which could facilitate the estimation of specific hydraulic parameters for the individual plants.

Our analysis of the transpiration beta factor that represents water stress in CLM5 revealed a lack of correlation between transpiration beta and the simulated annual yield. Reducing the precipitation in the forcing data sets generally had an effect on the transpiration beta factor, which reflects the induced water stress at lower precipitation rates. It appears that the threshold for plant water stress induced yield loss is not reached in the simulations with unchanged WFDE5 precipitation amounts or its effects on overall grain yield remain too small. The small range of transpiration water stress in simulation results implies that the modeled system experienced a relatively uniform degree of water stress, possibly due to low variations in soil moisture conditions or other factors influencing water availability for transpiration, such as the water-retaining capacities of the soil or irrigation. While the effects of irrigation can be considered negligible in our predominantly rain-fed simulation domains, this small range of beta serves as an indication for inconsistencies in the representation of the soil moisture regime. In addition to limitations in representing crop yield variability, we also observed profound differences of the simulated soil moisture contents throughout the decades compared to the ESA-CCI and the SMAP L3 products. For DE-NRW, the simulated surface soil moisture was systematically overestimated compared to both ESA-CCI and SMAP L3 during the early growing season. The same was observed for the AUS-VIC simulation results compared to ESA-CCI, where CLM5 simulations resulted in higher daily SMCs. Compared to SMAP L3, however, the simulated SMC fitted well during the early stages of the year, while daily values of SMAP L3 during the growing season were underestimated in the simulation results. The observed differences in error between the simulated soil moisture content (SMC) by CLM5 and satellite-

derived data (ESA-CCI and SMAP L3) at different growth stages can be attributed to several factors specific to each region. Overall, these discrepancies in error between simulated and satellite-derived SMC at different growth stages could stem from variations in soil properties, precipitation patterns, and uncertainties associated with satellite-derived data retrieval algorithms and spatial resolution, all of which interact differently in each region and influence soil moisture dynamics throughout the year. The discrepancy during winter months may be due to greater biases in satellite-derived products related to frozen soil and snow cover. Moreover, the higher SMC in CLM5 simulations outside of the main cropping season in DE-NRW could also result from a misrepresentation of post-harvest field conditions, where large fields of cropland are simulated as baer soil while, in reality, cover crops or weeds are growing on these fields. A misrepresentation of the soil moisture regime with overly high soil moisture contents could potentially dampen the potential benefits of the new plant hydraulic stress routine in CLM5. In earlier studies, data assimilation has been applied to address discrepancies between CLM simulated soil moisture and data derived from satellite and field observations (e.g., Hung et al., 2022; Naz et al., 2019; Strebel et al., 2022; Zhao et al., 2021). However, data assimilation of soil moisture and groundwater level observations had only limited effects on simulated evapotranspiration (Hung et al., 2022). Another way to improve the predictions of hydrologic states and fluxes with LSMs is the coupling with subsurface or groundwater models, such as ParFlow (Kollet & Maxwell, 2008; Kuffour et al., 2020), in integrated modeling approaches (e.g., Kollet & Maxwell, 2008; Maxwell et al., 2011; Naz et al., 2023; Soltani et al., 2022; Tian et al., 2012; Yuan et al., 2008). These studies have shown that coupled models can simulate complex processes more realistically than uncoupled models (e.g., Maxwell et al., 2011; Tian et al., 2012; Yuan et al., 2008). In addition, it is crucial to acknowledge the uncertainties in the satellite-derived data sets. Studies evaluating the quality of SMAP in Europe found local errors of 0.056 cm^3/cm^3 for a catchment in the state of NRW, Germany (Zhao et al., 2021), and comparable error magnitudes close to 0.06 cm^3/cm^3 for a region in the Netherlands (van der Velde et al., 2021). Addressing these uncertainties is crucial for interpreting satellite-derived soil moisture data accurately, with ongoing research and technological advancements contributing to improvements in reliability. For example, Seo and Dirmeyer (2022) propose an adjustment of ESA-CCI soil moisture achieved through Fourier transform time-filtering, which led to improved subseasonal variability, increased temporal correlation, and enhanced skill across various land cover classes.

Moreover, the simulation of soil water fluxes is intricately linked to the hydraulic properties of the soil, which are in CLM5 estimated from soil texture information using pedotransfer functions after Clapp and Hornberger (1978) or Cosby et al. (1984). In addition, the use of different pedotransfer functions for specific soil types can introduce substantial variability in the numerical modeling results of water fluxes, as demonstrated by Weihermüller et al. (2021) and Boas and Mallants (2022). One approach to improve the simulation of soil moisture in CLM5 may therefore involve the numerical implementation of various pedotransfer functions tailored to different soil types. Alternatively, the soil hydraulic information could be acquired through machine learning techniques or by utilizing suitable pedotransfer functions, similar to the approach outlined by Montzka et al. (2017), and incorporated via input files similar to the other surface input data. The inherent uncertainty in simulation results due to input data extends also to meteorological information. Bodjrenou et al. (2023) assessed multiple reanalysis products, including WFDE5, in West Africa from 1981 to 2019. While WFDE5 provided accurate estimates for annual precipitation, it exhibited an excess of small rainfall events compared to observations. These uncertainties in commonly used reanalysis products for land surface modeling add an additional layer of complexity when examining the inter-annual variability of simulated variables influenced by precipitation.

While the adequate simulation of plant water stress is crucial to simulate crop productivity in response to changing weather conditions, it is also important to address additional processes that influence the yield variability of rain-fed agriculture. Crop management practices (e.g., fertilization, double cropping) and the selection of cultivars specifically bred for high grain yields under local climate conditions have significant impacts on agricultural production and are currently not accounted for in CLM5. Allowing for a broader range of yields (maximum productivity) for specific crops through changes in the physiological parameters that constrain the simulation of crop growth and development in CLM5 could be another approach to allow for a higher inter-annual variability of yield. Whether structural modifications to the phenology module or systematic adjustments in the crop-specific parameterization can effectively allow for a broader range of yield magnitudes for the respective crops remains to be evaluated through model evaluation and parameter sensitivity studies.

In summary, we argue that the limitations of model performance presented and discussed in this study arise from three main sources. First, despite the incorporation of the new plant hydraulic stress routine in CLM5, there is a lack of sensitivity of crop yield towards drought stress and soil water availability. This may arise from the simplified parameterization of plant hydraulics which is unified for all plant functional types in CLM5. Second, CLM5 simulated soil moisture exhibited higher values during the cropping season, particularly in the early stage, than both ESA-CCI and SMAP products, which may have contributed to the weaker inter annual variability of yield. Third, while CLM5 already incorporates a complex representation of crop growth and phenology compared with other state-of-the-art LSMs, there are several significant factors that contribute to inter-annual yield variability that are currently not adequately accounted for. Some examples include human crop management practices in response to technological advancements and public policies, diverse farming techniques such as varying fertilization amounts and types, double-cropping, as well as environmental factors such as pests, diseases, and floods. Additionally, the crop–specific parameterizations in the crop phenology module, as well as the limited availability of data representing a wide range of crop varieties and geographic regions, represent substantial areas for improvement.

To overcome these challenges, the plant specific hydraulic parameterization needs to be improved, which can be achieved with the help of high-resolution field observations combined with parameter estimation methods. For a better representation of crop growth and yield, there is a need for further model developments, representing the influence of frost, pests, hail and wind on crop growth, different fertilizer types and application techniques and a more detailed representation of root crops. Finally, we propose to use state and parameter updating techniques, such as data assimilation, when studying global climate change impacts on agriculture with CLM5 to better account for model and parameter uncertainties (e.g., Strebel et al., 2022).

5. Summary and Conclusions

Rapid changes in atmospheric conditions and land use over recent decades have made the fate of our terrestrial biosphere, depletion of natural water resources, food security, and the impact of anthropogenic carbon emissions major global research topics. Land surface models (LSMs), such as CLM5, are the primary tools used to study changes in our terrestrial surface in response to climate projections. With predicted changes in regional and global climate, and a potential increase in drought risk, it is vital for the LSMs used in coupled climate models to realistically portray the drought responses of the land surface, and vegetation in particular. Reliable predictions of crop yield variability can contribute to discussions on climate change impacts and mitigation strategies.

In this study, we evaluated the performance of the land surface model CLM5 forced with reanalysis data in representing inter-annual variability of crop yield in multi-decadal simulations for two regions, AUS-VIC and DE-NRW, in different climate zones and with different soil moisture regimes. Evaluation studies for different ecosystems, such as the one presented here, are essential to improve our understanding of model performance and to identify the key challenges toward reliable projections of LSMs in climate change research.

Our analysis showed that CLM5 was able to capture the overall magnitudes of yields for individual crops, as well as regional differences for the same crops in the two domains (i.e., lower overall yield magnitudes for AUS-VIC than for DE-NRW for the same crop type). Overall, the lower annual yields per area over AUS-VIC can be attributed to differences in climate, crop varieties grown, soil characteristics, fertilizer rates and management techniques (e.g., larger paddock sizes with less dense plantations). Previous studies showed that the yield magnitudes for winter cereals specifically were too low at multiple European test sites (Boas et al., 2021). Hence, modifications and enhancements were introduced and validated for several European sites by Boas et al. (2021). These modifications were also used in our study, contributing to the reasonable yield magnitudes observed for DE-NRW. These factors collectively contribute to the disparities observed in simulation results between the two domains. However, the inter-annual fluctuations of yield in response to differences in weather patterns, such as seasonal rainfall amounts, were underestimated for both domains. Interestingly, the higher variability throughout the AUS-VIC domain compared to DE-NRW was reflected in the simulation results. Analysis of the plant water stress regime and correlations between seasonal rainfall amounts and crop yields revealed a misrepresentation of the more water-limited regime in AUS-VIC.

Experiments with reduced precipitation amounts that synthetically increased the plant water stress in simulations were able to better capture inter-annual variations of crop yield but underestimated crop yield. CLM5 is typically fine-tuned to replicate average crop yields through parameter adjustments. Theoretically, well-calibrated parameters have the potential to offset the impacts environmental factors such as pests, diseases, wildlife damage or pollutants.

Possible explanations for the underestimation of inter-annual variability of crop yields include: (a) a lack of sensitivity within the vegetation and crop module toward changes in soil moisture contents and soil water available for plants, possibly due to the parameterization of plant hydraulics; (b) systematic wet bias in simulated soil moisture content that may have dampened the potential benefit of the new plant water stress representation in CLM5; and (c) general uncertainties in the simulation of crop growth and yield due to inaccurate parameterizations and underrepresentation of environmental factors (e.g., pests, diseases) and human influences (e.g., agricultural management decisions, fertilizer type and application, cultivar selection). To remedy those effects, model enhancements are necessary, particularly in the field of plant hydraulic and physiological parameterizations. In addition, integrated approaches with subsurface or groundwater models (e.g., ParFlow) along with data assimilation offer potential for enhancing the simulation accuracy of soil moisture and other variables in future studies.

Appendix A

Journal of Advances in Modeling Earth Systems

10.1029/2023MS004023

Table A1

Total Annual Crop Yield (t/ha) and Corresponding Yield Anomaly (%) From Simulation Results (CLM) for Winter Wheat, Winter Wheat Monoculture With Unchanged WFDE5 Precipitation (CLM-WFDE5) and Reduced Precipitation (CLM_LowP), Barley, Canola and Sorghum Throughout the AUS-VIC Domain for the Years 1999–2019, Compared to Available Records (Obs) From ABARES (2022)

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Yield [t/ha]																					
Obs																					
Winter wheat	2.14	2.70	2.46	0.72	2.23	1.45	2.21	0.65	1.32	1.14	1.66	2.46	2.36	2.15	2.21	1.76	1.35	3.21	2.54	1.62	2.60
Barley	2.03	2.41	2.36	0.62	2.61	1.41	2.31	0.66	1.62	1.29	1.91	2.43	2.41	2.29	2.22	1.50	1.31	3.26	2.50	1.50	2.83
Canola	1.39	1.46	1.46	0.71	1.61	1.21	1.43	0.24	1.13	0.82	1.43	1.47	1.44	1.47	1.62	1.16	1.04	1.94	1.73	1.23	1.81
Sorghum	7.32	1.88	-	-	0.43	1.32	2.64	1.91	1.92	-	-	2.01	-	0.56	1.27	-	2.78	0.34	0.61	0.72	-
Mean	1.85	2.19	2.09	0.68	2.15	1.36	1.98	0.52	1.35	1.08	1.67	2.12	2.07	1.97	2.01	1.47	1.23	2.80	2.26	1.45	2.41
CLM																					
Winter wheat	2.13	2.38	2.33	2.24	2.53	2.37	2.35	2.06	2.16	2.18	2.07	2.53	2.22	2.31	1.95	2.33	2.03	2.21	2.27	2.24	2.33
Barley	1.66	1.40	0.87	1.46	1.40	1.81	1.13	1.42	1.62	1.12	0.81	1.27	1.43	1.47	2.05	1.55	1.15	1.54	0.90	0.35	1.59
Canola	1.26	0.85	0.80	1.34	1.29	1.22	0.58	0.91	1.74	1.21	1.42	0.99	1.07	0.89	1.92	1.40	0.86	1.15	1.10	0.46	1.48
Sorghum	2.39	2.31	1.78	1.90	2.45	1.63	2.19	2.07	2.05	2.25	2.33	2.50	2.46	2.42	2.04	2.05	2.40	2.18	2.55	2.40	2.22
Mean	1.68	1.54	1.33	1.68	1.74	1.80	1.35	1.46	1.84	1.50	1.43	1.59	1.57	1.56	1.98	1.76	1.35	1.63	1.42	1.02	1.80
Winter wheat monoculture																					
CLM_WFDE5	1.81	2.13	2.20	2.13	2.32	2.06	2.13	1.90	1.90	1.79	1.67	2.07	1.89	1.98	1.68	1.96	1.59	1.68	1.86	1.82	1.76
CLM_LowP	1.39	1.65	1.70	1.46	1.54	1.36	1.51	0.90	1.34	0.97	1.32	1.71	1.48	1.24	1.33	1.02	0.86	1.37	1.38	1.17	1.30
Yield anomaly [%]																					
Obs																					
Winter wheat	9.72	38.18	25.93	−63.16	14.44	−25.55	13.44	−66.54	−32.43	−41.31	−14.75	26.16	21.13	10.23	13.36	−9.62	−30.64	64.46	30.48	−16.80	33.27
Barley	2.91	22.04	19.77	−68.84	32.15	−28.49	17.10	−66.47	−18.14	−34.86	−3.29	22.87	22.16	15.75	12.23	−24.05	−33.59	65.01	26.64	−24.14	43.26
Canola	5.25	10.16	10.14	−46.09	21.38	−8.88	7.92	−82.20	−14.80	−38.07	8.40	11.31	9.06	11.15	22.18	−12.59	−21.45	46.47	30.86	−6.82	36.61
Sorghum	298.88	2.25	-	-	−76.59	−27.98	43.82	4.14	4.41	-	-	9.35	-	−69.48	−31.07	-	51.60	−81.67	−66.66	−61.01	-
Mean	6.03	25.04	19.63	−60.99	22.85	−22.46	13.43	−70.46	−22.61	−38.07	−4.60	21.18	18.48	12.54	15.16	−15.80	−29.44	60.13	29.13	−17.05	37.87
CLM																					
Winter wheat	−5.34	5.99	3.58	−0.20	12.47	5.53	4.34	−8.46	−4.01	−3.27	−7.88	12.28	−1.22	2.86	−13.10	3.64	−9.93	−1.66	0.84	−0.20	3.75
Barley	26.55	6.73	−33.71	11.01	6.87	37.71	−13.76	8.43	23.28	−14.80	−38.31	−3.52	8.73	11.97	56.44	17.80	−12.18	17.44	−31.23	−73.38	21.21
Canola	9.76	−25.96	−30.04	16.56	12.23	6.51	−49.76	−20.72	51.37	5.48	23.73	−14.13	−7.17	−22.57	67.17	22.19	−25.27	0.42	−4.58	−60.27	28.84
Sorghum	8.48	4.85	−19.29	−13.95	11.01	−25.97	−0.44	−6.24	−6.96	2.17	5.67	13.39	11.61	9.97	−7.52	−7.02	8.94	−0.95	15.76	8.99	0.82
Mean	6.95	−1.84	−15.22	6.74	10.57	14.44	−14.11	−6.98	16.80	−4.59	−8.88	1.19	−0.15	−1.05	25.52	11.82	−14.51	3.90	−9.64	−35.39	14.44
Winter wheat monoculture																					
CLM_WFDE5	−5.65	11.04	14.35	10.97	20.92	7.19	10.72	−1.28	−1.14	−6.61	−13.23	7.70	−1.42	3.35	−12.47	2.06	−17.38	−12.37	−3.23	−4.95	−8.58
CLM_LowP	3.98	23.52	27.64	9.34	15.71	1.73	13.60	−32.81	0.75	−26.92	−1.12	27.93	11.07	−6.96	0.11	−23.62	−35.70	2.74	3.70	−12.26	−2.44

Journal of Advances in Modeling Earth Systems

10.1029/2023MS004023

Table A2

Total Annual Crop Yield (t/ha) and Corresponding Yield Anomaly (%) From Simulation Results (CLM) for Winter Wheat, Winter Wheat Monoculture With Unchanged WFDE5 Precipitation (CLM-WFDE5) and Reduced Precipitation (CLM_LowP), Spring Wheat, Canola and Corn Throughout the DE-NRW Domain for the Years 1999–2019, Compared to Available Records (Obs) From IT.NRW (2022)

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Yield [t/ha]																					
Obs																					
Winter wheat	-	-	-	-	-	-	8.55	7.83	7.29	8.93	8.65	7.79	8.13	8.44	9.15	9.08	8.81	8.07	7.94	7.98	8.17
Spring wheat	-	-	-	-	-	-	6.54	6.08	5.73	6.84	6.65	6.01	5.83	7.14	7.15	6.60	7.08	6.21	5.85	5.40	5.36
Canola	-	-	-	-	-	-	3.83	3.82	3.50	3.65	4.25	4.01	3.64	3.90	4.14	4.29	4.03	3.85	3.90	3.48	3.69
Corn	-	-	-	-	-	-	10.10	8.70	9.61	10.58	10.33	9.46	11.18	11.37	10.17	11.17	9.99	9.86	10.74	7.80	8.44
Mean	-	-	-	-	-	-	7.26	6.61	6.53	7.50	7.47	6.82	7.20	7.71	7.65	7.79	7.48	7.00	7.11	6.17	6.42
CLM																					
Winter wheat	8.58	7.20	8.11	7.49	7.73	8.26	7.92	7.62	7.46	8.30	7.91	8.03	7.98	8.03	7.71	8.20	7.72	8.13	7.79	7.60	8.23
Spring wheat	5.76	4.12	5.02	3.68	5.50	5.03	5.34	5.09	4.30	4.90	5.76	5.37	5.43	5.05	4.65	5.36	4.93	4.61	5.67	5.34	5.89
Canola	5.71	3.86	4.93	3.71	5.54	5.08	5.32	5.03	3.89	4.90	5.71	5.38	5.37	5.18	4.62	5.37	4.91	4.59	5.65	5.80	5.80
Corn	9.31	7.93	9.25	8.42	9.52	9.25	9.08	8.74	8.21	8.57	9.21	8.04	8.93	9.15	8.11	9.05	8.62	8.67	8.83	8.74	8.33
Mean	7.34	5.78	6.83	5.82	7.07	6.91	6.91	6.62	5.96	6.67	7.15	6.70	6.93	6.86	6.27	6.99	6.54	6.50	6.98	6.87	7.06
Winter wheat monoculture																					
CLM_WFDE5	8.56	7.20	8.13	7.49	7.74	8.29	7.92	7.64	7.46	7.97	7.91	8.08	7.99	8.03	7.72	8.10	7.73	8.13	7.80	7.60	8.24
CLM_LowP	8.58	7.08	8.17	7.26	7.14	7.98	7.15	5.19	7.23	7.19	7.12	6.25	6.06	7.71	7.42	7.52	6.87	7.75	6.83	6.41	6.25
Yield anomaly [%]																					
Obs																					
Winter wheat	-	-	-	-	-	-	2.76	−5.90	−12.39	7.32	3.96	−6.38	−2.29	1.43	9.97	9.13	5.88	−3.01	−4.57	−4.09	−1.81
Spring wheat	-	-	-	-	-	-	3.84	−3.46	−9.02	8.61	5.59	−4.57	−7.43	13.37	13.53	4.80	12.42	−1.40	−7.11	−14.26	−14.89
Canola	-	-	-	-	-	-	−0.91	−1.17	−9.45	−5.57	9.95	3.74	−5.83	0.90	7.11	10.99	4.26	−0.40	0.90	−9.97	−4.54
Corn	-	-	-	-	-	-	1.34	−12.71	−3.58	6.15	3.65	−5.08	12.17	14.08	2.04	12.07	0.23	−1.07	7.76	−21.74	−15.32
Mean	-	-	-	-	-	-	2.00	−7.10	−8.16	5.45	5.02	−4.15	1.16	8.43	7.59	9.45	5.13	−1.62	−0.07	−13.32	−9.81
CLM																					
Winter wheat	8.58	−8.92	2.54	−5.27	−2.19	4.47	0.17	−3.54	−5.68	4.97	0.10	1.56	1.02	1.59	−2.45	3.73	−2.31	2.80	−1.51	−3.81	4.14
Spring wheat	13.22	−18.91	−1.21	−27.58	8.11	−1.09	4.92	0.11	−15.43	−3.62	13.20	5.60	6.84	−0.59	−8.60	5.32	−3.13	−9.44	11.47	5.06	15.76
Canola	12.63	−23.74	−2.64	−26.76	9.35	0.30	5.08	−0.62	−23.14	−3.20	12.81	6.15	6.03	2.36	−8.80	5.94	−3.16	−9.33	11.58	14.53	14.60
Corn	6.28	−9.51	5.62	−3.93	8.72	5.62	3.67	−0.20	−6.33	−2.21	5.09	−8.21	1.99	4.47	−7.43	3.31	−1.58	−1.01	0.81	−0.25	−4.92
Mean	9.47	−13.81	1.86	−13.12	5.50	3.00	3.14	−1.20	−11.04	−0.55	6.62	0.00	3.38	2.26	−6.44	4.31	−2.39	−3.06	4.18	2.50	5.36
Winter wheat monoculture																					
CLM_WFDE5	8.51	−8.73	3.02	−5.10	−1.90	5.05	0.32	−3.22	−5.51	1.05	0.27	2.43	1.23	1.70	−2.22	2.65	−2.11	2.99	−1.20	−3.65	4.41
CLM_LowP	20.76	−0.26	15.06	2.20	0.60	12.29	0.74	−26.89	1.73	1.19	0.22	−12.01	−14.67	8.53	4.41	5.81	−3.26	9.09	−3.84	−9.70	−11.99

Table A3

Statistical Analysis Results R^2, t-Statistics and Probability Values (p-Values) for Simulated Annual Mean Crop Yield (Averaged for All Regarded Crops) in Relation to the Evaluated Variables Seasonal Mean Root Zone Soil Moisture (0.02–0.32 m Depth) [m^3/m^3], Seasonal WFDE5 Rainfall Amounts [mm], and Transpiration Beta [-], for the AUS-VIC and DE-NRW Domain, Respectively

AUS-VIC	R^2	t-stat	*p*-value
Root zone soil moisture	0.07	1.15	0.26
Seasonal rainfall	0.18	2.06	0.05
Transpiration beta	0.00	0.01	0.99
DE_NRW	R^2	t-stat	*p*-value
Root zone soil moisture	0.21	2.22	0.04
Seasonal rainfall	0.21	2.28	0.03
Transpiration beta	0.18	2.05	0.05

Note. Additional Parameters can be Found in Supporting Information S1

Table A4

Multiple Correlation Coefficient (Multiple r), R^2, Adjusted R^2 (R^2 Adjusted for the Complexity of the Model), Standard Error and Corresponding t-statistics and Probability Values (p-Values) Resulting From Multiple Regression Analysis for Simulated Annual Mean Crop Yield (Averaged for All Regarded Crops, Dependent Variable), Explained With the Simulated Mean Seasonal Root Zone Soil Moisture and the Mean Seasonal WFDE5 Global Radiation as Independent Variables, and With the Seasonal WFDE5 Precipitation Amount and Mean Seasonal WFDE5 Global Radiation as Independent Variables, for the DE-NRW and AUS-VIC Domain, Respectively

Annual mean crop yield - mean seasonal root zone soil moisture and mean seasonal global radiation						
DE-NRW						
Multiple *r*	0.6146		Coefficients	Standard error	t-statistics	*p*-value
R^2	0.3777	Intercept	4.5111	2.7914	1.6161	0.1235
Adjusted R^2	0.3086	Root zone soil moisture	−6.9219	4.5068	−1.5359	0.1420
Standard error	0.3563	Global radiation	0.0356	0.0159	2.2312	0.0386
AUS-VIC						
Multiple r	0.4081		Coefficients	Standard error	t-statistics	*P*-value
R^2	0.1665	Intercept	−3.3872	2.9009	−1.1676	0.2582
Adjusted R^2	0.0739	Root zone soil moisture	5.5062	3.0850	1.7849	0.0911
Standard error	0.2102	Global radiation	0.0182	0.0123	1.4782	0.1566
Annual mean crop yield - seasonal precipitation amount and mean seasonal global radiation						
DE-NRW						
Multiple r	0.5945		Coefficients	Standard error	t-statistics	*P*-value
R^2	0.3534	Intercept	3.1432	2.3843	1.3183	0.2039
Adjusted R^2	0.2816	Seasonal rainfall	−0.0016	0.0013	−1.2625	0.2229
Standard error	0.3631	Global radiation	0.0339	0.0172	1.9677	0.0647
AUS-VIC						
Multiple r	0.5864		Coefficients	Standard error	t-statistics	*P*-value
R^2	0.3439	Intercept	−3.6735	2.2545	−1.6294	0.1206
Adjusted R^2	0.2710	Seasonal rainfall	0.0034	0.0012	2.9854	0.0079
Standard error	0.1865	Global radiation	0.0226	0.0107	2.1063	0.0495

Data Availability Statement

The modified model version of CLM5 that was used in this study is publicly accessible as a supplement to Boas et al. (2021). All underlying research data used for this study is publicly accessible. The WFDE5 data set is distributed by the Copernicus Climate Change Service (C3S) through its Climate Data Store (CDS) (Copernicus Climate Change Service, 2018). Soil information from SoilGrids are publicly accessible via the International Soil Reference and Information Center (ISRIC) - World Soil Information data hub (ISRIC, 2023). Land cover information that was used for the DE-NRW domain can be accessed via https://doi.pangaea.de/10.1594/PANGAEA.893195 (Griffiths et al., 2018). Land cover information for AUS-VIC from the Victorian Land Use Information System is publicly accessible via the Victoria Government Data Directory (Victoria Government Data Directory, 2018). The crop yield statistics from the Australian Crop Report can be accessed at https://doi.org/10.25814/xqy3-sx57 and https://doi.org/10.25814/0c4s-qd09 (ABARES, 2020, 2022) and the yield records for NRW are available at https://www.it.nrw/statistik/eckdaten/erntevon-ausgewaehlten-landwirtschaftlichen-feldfruechten-und-gruenland-767 (IT.NRW, 2022). Soil moisture products used in this study are available at https://nsidc.org/data/spl3smap/versions/3 (Entekhabi et al., 2016) for SMAP L3 and ESA-CCI at https://doi.pangaea.de/10.1594/PANGAEA.940409 (Hongtao et al., 2022).

Acknowledgments

The authors gratefully acknowledge the Earth System Modeling Project (ESM) for funding this work by providing computing time on the ESM partition of the supercomputer JUWELS (Jülich Supercomputing Center. JUWELS Cluster and Booster: Exascale Pathfinder with Modular Supercomputing Architecture at Juelich Supercomputing Center. Journal of large-scale research facilities, 7, A183. http://dx.doi.org/10.17815/jlsrf-7-183, 2021) at the Jülich Supercomputing Center (JSC). This study is part of the Juelich-University of Melbourne Postgraduate Academy (JUMPA), an international research collaboration between the University of Melbourne, Australia, and the Research Center Juelich, Germany. This work was also supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2070–390732324, project PhenoRob. The authors gratefully acknowledge the editor and the reviewers for their helpful feedback and comments, which have contributed to the improvement of this manuscript. Open Access funding enabled and organized by Projekt DEAL.

References

ABARES. (2020). Australian Bureau of agricultural and resource economics and Sciences: Australian crop report, February 2021 [Dataset]. *Canberra*. https://doi.org/10.25814/xqy3-sx57

ABARES. (2022). Australian Bureau of agricultural and resource economics and Sciences: Australian crop report, December 2022 [Dataset]. *Canberra*. https://doi.org/10.25814/0c4s-qd09

Asseng, S., Martre, P., Maiorano, A., Rötter, R. P., O'Leary, G. J., Fitzgerald, G. J., et al. (2019). Climate change impact and adaptation for wheat protein. *Global Change Biology*, *25*(1), 155–173. https://doi.org/10.1111/gcb.14481

Blanchard, J. L., Watson, R. A., Fulton, E. A., Cottrell, R. S., Nash, K. L., Bryndum-Buchholz, A., et al. (2017). Linked sustainability challenges and trade-offs among fisheries, aquaculture and agriculture. *Nature Ecology and Evolution*, *1*(9), 1240–1249. https://doi.org/10.1038/s41559-017-0258-8

Blyth, E. M., Arora, V. K., Clark, D. B., Dadson, S. J., De Kauwe, M. G., Lawrence, D. M., et al. (2021). Advances in land surface modelling. *Current Climate Change Reports*, *7*(2), 45–71. https://doi.org/10.1007/s40641-021-00171-5

BMEL. (2020). Besondere Ernte-und Qualitätsermittlung (BEE) 2019, agricultural yield and quality assessment. [Dataset]. https://www.bmel-statistik.de/landwirtschaft/ernte-und-qualitaet/archiv-ernte-und-qualitaet-bee

BMEL. (2022). Besondere Ernte-und Qualitätsermittlung (BEE) 2021, agricultural yield and quality assessment. [Dataset]. https://www.bmel-statistik.de/landwirtschaft/ernte-und-qualitaet/archiv-ernte-und-qualitaet-bee

Boas, T., Bogena, H., Grünwald, T., Heinesch, B., Ryu, D., Schmidt, M., et al. (2021). Improving the representation of cropland sites in the community land model (CLM) version 5.0. *Geoscientific Model Development*, *14*(1), 573–601. https://doi.org/10.5194/gmd-14-573-2021

Boas, T., Bogena, H., Ryu, D., Vereecken, H., Western, A., & Hendricks-Franssen, H.-J. (2023). Seasonal crop yield prediction with SEAS5 long-range meteorological forecasts in a land surface modelling approach. *Hydrology and Earth System Sciences Discussions*, 1–30. https://doi.org/10.5194/hess-2023-28

Boas, T., & Mallants, D. (2022). Episodic extreme rainfall events drive groundwater recharge in arid zone environments of central Australia. *Journal of Hydrology: Regional Studies*, *40*, 101005. https://doi.org/10.1016/j.ejrh.2022.101005

Bodjrenou, R., Cohard, J.-M., Hector, B., Lawin, E. A., Chagnaud, G., Danso, D. K., et al. (2023). Evaluation of reanalysis estimates of precipitation, radiation, and temperature over Benin (West Africa). *Journal of Applied Meteorology and Climatology*, *62*(8), 1005–1022. https://doi.org/10.1175/JAMC-D-21-0222.1

Cammarano, D., Valdivia, R. O., Beletse, Y. G., Durand, W., Crespo, O., Tesfuhuney, W. A., et al. (2020). Integrated assessment of climate change impacts on crop productivity and income of commercial maize farms in northeast South Africa. *Food Security*, *12*(3), 659–678. https://doi.org/10.1007/s12571-020-01023-0

Challinor, A. J., Watson, J., Lobell, D. B., Howden, S. M., Smith, D. R., & Chhetri, N. (2014). A meta-analysis of crop yield under climate change and adaptation. *Nature Climate Change*, *4*, 287–291. https://doi.org/10.1038/nclimate2153

Clapp, R. B., & Hornberger, G. M. (1978). Empirical equations for some soil hydraulic properties. *Water Resources Research*, *14*(4), 601–604. https://doi.org/10.1029/WR014i004p00601

Claverie, M., Ju, J., Masek, J. G., Dungan, J. L., Vermote, E. F., Roger, J.-C., et al. (2018). The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sensing of Environment*, *219*, 145–161. https://doi.org/10.1016/j.rse.2018.09.002

Collatz, G. J., Ribas-Carbo, M., & Berry, J. A. (1992). Coupled photosynthesis-stomatal conductance model for leaves of C4 plants. *Functional Plant Biology*, *19*(5), 519–538. https://doi.org/10.1071/pp9920519

Copernicus Climate Change Service. (2018). Climate Data Store: Seasonal forecast daily and subdaily data on single levels [Dataset]. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. https://doi.org/10.24381/cds.181d637e

Cosby, B. J., Hornberger, G. M., Clapp, R. B., & Ginn, T. R. (1984). A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resources Research*, *20*(6), 682–690. https://doi.org/10.1029/WR020i006p00682

Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., et al. (2020). WFDE5: Bias-adjusted ERA5 reanalysis data for impact studies. *Earth System Science Data*, *12*(3), 2097–2120. https://doi.org/10.5194/essd-12-2097-2020

Dagon, K., Sanderson, B. M., Fisher, R. A., & Lawrence, D. M. (2020). A machine learning approach to emulation and biophysical parameter estimation with the Community Land Model, version 5. *Advances in Statistical Climatology, Meteorology and Oceanography*, *6*(2), 223–244. https://doi.org/10.5194/ascmo-6-223-2020

De Kauwe, M. G., Kala, J., Lin, Y.-S., Pitman, A. J., Medlyn, B. E., Duursma, R. A., et al. (2015a). A test of an optimal stomatal conductance scheme within the CABLE land surface model. *Geoscientific Model Development*, *8*(2), 431–452. https://doi.org/10.5194/gmd-8-431-2015

De Kauwe, M. G., Zhou, S.-X., Medlyn, B. E., Pitman, A. J., Wang, Y.-P., Duursma, R. A., & Prentice, I. C. (2015). Do land surface models need to include differential plant species responses to drought? Examining model predictions across a mesic-xeric gradient in Europe. *Biogeosciences*, *12*(24), 7503–7518. https://doi.org/10.5194/bg-12-7503-2015

Denissen, J. M. C., Teuling, A. J., Pitman, A. J., Koirala, S., Migliavacca, M., Li, W., et al. (2022). Widespread shift from ecosystem energy to water limitation with climate change. *Nature Climate Change*, *12*(7), 677–684. https://doi.org/10.1038/s41558-022-01403-8

Deryng, D., Conway, D., Ramankutty, N., Price, J., & Warren, R. (2014). Global crop yield response to extreme heat stress under multiple climate change futures. *Environmental Research Letters*, *9*(3), 034011. https://doi.org/10.1088/1748-9326/9/3/034011

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., et al. (2017). ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment*, *203*, 185–215. https://doi.org/10.1016/j.rse.2017.07.001

Entekhabi, D., Das, N., Njoku, E., Johnson, J., & Shi, J. (2016). SMAP L3 radar/radiometer global daily 9 km EASE-grid soil moisture, version 3 [Dataset]. https://doi.org/10.5067/7KKNQ5UURM2W

Farquhar, G. D., von Caemmerer, S., & Berry, J. A. (1980). A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. *Planta*, *149*(1), 78–90. https://doi.org/10.1007/BF00386231

Fisher, R. A., & Koven, C. D. (2020). Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems. *Journal of Advances in Modeling Earth Systems*, *12*(4), e2018MS001453. https://doi.org/10.1029/2018MS001453

Franke, J. A., Müller, C., Elliott, J., Ruane, A. C., Jägermeyr, J., Balkovic, J., et al. (2020). The GGCMI phase 2 experiment: Global gridded crop model simulations under uniform changes in CO2, temperature, water, and nitrogen levels (protocol version 1.0). *Geoscientific Model Development*, *13*(5), 2315–2336. https://doi.org/10.5194/gmd-13-2315-2020

Franks, P. J., Bonan, G. B., Berry, J. A., Lombardozzi, D. L., Holbrook, N. M., Herold, N., & Oleson, K. W. (2018). Comparing optimal and empirical stomatal conductance models for application in Earth system models. *Global Change Biology*, *24*(12), 5708–5723. https://doi.org/10.1111/gcb.14445

French, R. J., & Schultz, J. E. (1984a). Water use efficiency of wheat in a Mediterranean-type environment. I. The relation between yield, water use and climate. *Australian Journal of Agricultural Research*, *35*(6), 743–764. https://doi.org/10.1071/ar9840743

French, R. J., & Schultz, J. E. (1984b). Water use efficiency of wheat in a Mediterranean-type environment. II. some limitations to efficiency. *Australian Journal of Agricultural Research*, *35*(6), 765–775. https://doi.org/10.1071/ar9840765

Griffiths, P., Nendel, C., & Hostert, P. (2018). National-scale crop- and land-cover map of Germany (2016) based on imagery acquired by Sentinel-2A MSI and Landsat-8 OLI. [Dataset]. PANGAEA. https://doi.org/10.1594/PANGAEA.893195

Griffiths, P., Nendel, C., & Hostert, P. (2019). Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. *Remote Sensing of Environment*, *220*, 135–151. https://doi.org/10.1016/j.rse.2018.10.031

Gruber, A., Dorigo, W. A., Crow, W., & Wagner, W. (2017). Triple collocation-based merging of satellite soil moisture retrievals. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(12), 6780–6792. https://doi.org/10.1109/TGRS.2017.2734070

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

Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations – The CRU TS3.10 dataset. *International Journal of Climatology*, *34*(3), 623–642. https://doi.org/10.1002/joc.3711

Hengl, T., de Jesus, J. M., Heuvelink, G. B. M., Gonzalez, M. R., Kilibarda, M., Blagotić, A., et al. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One*, *12*(2), e0169748. https://doi.org/10.1371/journal.pone.0169748

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hongtao, J., Huanfeng, S., Xinghua, L., & Lili, L. (2022). The 43-year (1978-2020) global 9km remotely sensed soil moisture product [Dataset]. *PANGAEA*. https://doi.org/10.1594/PANGAEA.940409

Hung, C. P., Schalge, B., Baroni, G., Vereecken, H., & Hendricks Franssen, H.-J. (2022). Assimilation of groundwater level and soil moisture data in an integrated land surface-subsurface model for southwestern Germany. *Water Resources Research*, *58*(6), e2021WR031549. https://doi.org/10.1029/2021WR031549

Huntzinger, D. N., Schwalm, C., Michalak, A. M., Schaefer, K., King, A. W., Wei, Y., et al. (2013). The North American carbon program multi-scale synthesis and terrestrial model intercomparison project – Part 1: Overview and experimental design. *Geoscientific Model Development*, *6*, 2121–2133. https://doi.org/10.5194/gmd-6-2121-2013

ISRIC. (2023). World soil information data hub: SoilGrids. [Dataset]. International Soil Reference and Information Centre (ISRIC). https://www.isric.org/explore/soilgrids

IT.NRW. (2022). Ernte ausgewählter landwirtschaftlicher Feldfrüchte, yield statistics for certain cash crops [Dataset]. *Landesbetrieb ITNRW, Düsseldorf*. https://www.it.nrw/statistik/eckdaten/erntevon-ausgewaehlten-landwirtschaftlichen-feldfruechten-und-gruenland-767

Jägermeyr, J., Müller, C., Ruane, A. C., Elliott, J., Balkovic, J., Castillo, O., et al. (2021). Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. *Nature Food*, *2*(11), 873–885. https://doi.org/10.1038/s43016-021-00400-y

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin America Meteorology Social*, *77*(3), 437–447. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Kennedy, D., Swenson, S., Oleson, K. W., Lawrence, D. M., Fisher, R., da Costa, A. C. L., & Gentine, P. (2019). Implementing plant hydraulics in the community land model, version 5. *Journal of Advances in Modeling Earth Systems*, *11*(2), 485–513. https://doi.org/10.1029/2018MS001500

Kimball, B. A., Boote, K. J., Hatfield, J. L., Ahuja, L. R., Stockle, C., Archontoulis, S., et al. (2019). Simulation of maize evapotranspiration: An inter-comparison among 29 maize models. *Agricultural and Forest Meteorology*, *271*, 264–284. https://doi.org/10.1016/j.agrformet.2019.02.037

Kollet, S. J., & Maxwell, R. M. (2008). Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model. *Water Resources Research*, *44*(2). https://doi.org/10.1029/2007WR006004

Konings, A. G., Yu, Y., Xu, L., Yang, Y., Schimel, D. S., & Saatchi, S. S. (2017). Active microwave observations of diurnal and seasonal variations of canopy water content across the humid African tropical forests. *Geophysical Research Letters*, *44*(5), 2290–2299. https://doi.org/10.1002/2016GL072388

Koster, R. D., Schubert, S. D., & Suarez, M. J. (2009). Analyzing the concurrence of meteorological droughts and warm periods, with implications for the determination of evaporative regime. *Journal of Climate*, *22*(12), 3331–3341. https://doi.org/10.1175/2008JCLI2718.1

Kuffour, B. N. O., Engdahl, N. B., Woodward, C. S., Condon, L. E., Kollet, S., & Maxwell, R. M. (2020). Simulating coupled surface–subsurface flows with ParFlow v3.5.0: Capabilities, applications, and ongoing development of an open-source, massively parallel, integrated hydrologic model. *Geoscientific Model Development*, *13*(3), 1373–1397. https://doi.org/10.5194/gmd-13-1373-2020

Lawrence, D. M., Fisher, R., Koven, C., Oleson, K., Svenson, S., Vertenstein, M., et al. (2018). *Technical description of version 5.0 of the community land model (CLM)*. National Center for Atmospheric Research (NCAR). http://www.cesm.ucar.edu/models/cesm2/land/CLM50_Tech_Note.pdf

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., et al. (2019). The community land model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245–4287. https://doi.org/10.1029/2018MS001583

Lawrence, D. M., Hurtt, G. C., Arneth, A., Brovkin, V., Calvin, K. V., Jones, A. D., et al. (2016). The land use model intercomparison project (LUMIP) contribution to CMIP6: Rationale and experimental design. *Geoscientific Model Development*, *9*, 2973–2998. https://doi.org/10.5194/gmd-9-2973-2016

Levis, S., Badger, A., Drewniak, B., Nevison, C., & Ren, X. (2018). CLMcrop yields and water requirements: Avoided impacts by choosing RCP 4.5 over 8.5. *Climate Change*, *146*(3–4), 501–515. https://doi.org/10.1007/s10584-016-1654-9

Li, L., Wang, Y.-P., Yu, Q., Pak, B., Eamus, D., Yan, J., et al. (2012). Improving the responses of the Australian community land surface model (CABLE) to seasonal drought. *Journal of Geophysical Research*, *117*(G4). https://doi.org/10.1029/2012JG002038

Li, Y., Guan, K., Gentine, P., Konings, A. G., Meinzer, F. C., Kimball, J. S., et al. (2017). Estimating global ecosystem Isohydry/Anisohydry using active and passive microwave satellite data. *Journal of Geophysical Research: Biogeosciences*, *122*(12), 3306–3321. https://doi.org/10.1002/2017JG003958

Lin, Y.-S., Medlyn, B. E., Duursma, R. A., Prentice, I. C., Wang, H., Baig, S., et al. (2015). Optimal stomatal behaviour around the world. *Nature Climate Change*, *5*, 459–464. https://doi.org/10.1038/nclimate2550

Lobell, D. B., Bala, G., & Duffy, P. B. (2006). Biogeophysical impacts of cropland management changes on climate. *Geophysical Research Letters*, *33*(6). https://doi.org/10.1029/2005GL025492

Lombardozzi, D. L., Lu, Y., Lawrence, P. J., Lawrence, D. M., Swenson, S., Oleson, K. W., et al. (2020). Simulating agriculture in the community land model version 5. *Journal of Geophysical Research: Biogeosciences*, *125*(8), e2019JG005529. https://doi.org/10.1029/2019JG005529

Lu, Y., Williams, I. N., Bagley, J. E., Torn, M. S., & Kueppers, L. M. (2017). Representing winter wheat in the community land model (version 4.5). *Geoscientific Model Development*, *10*(5), 1873–1888. https://doi.org/10.5194/gmd-10-1873-2017

Maxwell, R. M., Lundquist, J. K., Mirocha, J. D., Smith, S. G., Woodward, C. S., & Tompson, A. F. B. (2011). Development of a coupled groundwater–atmosphere model. *Monthly Weather Review*, *139*(1), 96–116. https://doi.org/10.1175/2010MWR3392.1

McDermid, S., Nocco, M., Lawston-Parker, P., Keune, J., Pokhrel, Y., Jain, M., et al. (2023). Irrigation in the Earth system. *Nature Reviews Earth & Environment*, *4*(7), 1–19. https://doi.org/10.1038/s43017-023-00438-5

Medlyn, B. E., Duursma, R. A., Eamus, D., Ellsworth, D. S., Prentice, I. C., Barton, C. V. M., et al. (2011). Reconciling the optimal and empirical approaches to modelling stomatal conductance. *Global Change Biology*, *17*(6), 2134–2144. https://doi.org/10.1111/j.1365-2486.2010.02375.x

Montzka, C., Herbst, M., Weihermüller, L., Verhoef, A., & Vereecken, H. (2017). A global data set of soil hydraulic properties and sub-grid variability of soil water retention and hydraulic conductivity curves. *Earth System Science Data*, *9*(2), 529–543. https://doi.org/10.5194/essd-9-529-2017

Morse-McNabb, E., Sheffield, K., Clark, R., Lewis, H., Robson, S., Cherry, D., & Williams, S. (2015). VLUIS, a land use data product for Victoria, Australia, covering 2006 to 2013. *Scientific Data*, *2*(1), 150070. https://doi.org/10.1038/sdata.2015.70

Müller, C., Elliott, J., Chryssanthacopoulos, J., Deryng, D., Folberth, C., Pugh, T. A. M., & Schmid, E. (2015). Implications of climate mitigation for future agricultural production. *Environmental Research Letters*, *10*(12), 125004. https://doi.org/10.1088/1748-9326/10/12/125004

Naz, B. S., Kurtz, W., Montzka, C., Sharples, W., Goergen, K., Keune, J., et al. (2019). Improving soil moisture and runoff simulations at 3 km over Europe using land surface data assimilation. *Hydrology and Earth System Sciences*, *23*(1), 277–301. https://doi.org/10.5194/hess-23-277-2019

Naz, B. S., Sharples, W., Ma, Y., Goergen, K., & Kollet, S. (2023). Continental-scale evaluation of a fully distributed coupled land surface and groundwater model, ParFlow-CLM (v3.6.0), over Europe. *Geoscientific Model Development*, *16*(6), 1617–1639. https://doi.org/10.5194/gmd-16-1617-2023

Nemani, R. R., Keeling, C. D., Hashimoto, H., Jolly, W. M., Piper, S. C., Tucker, C. J., et al. (2003). Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science*, *300*(5625), 1560–1563. https://doi.org/10.1126/science.1082750

NRW (North Rhine-Westphalia) State Government. (2020). Preliminary data on cereal grain harvest balance in 2020. Retrieved from https://www.land.nrw/pressemitteilung/nordrhein-westfalen-legt-erntebilanz-2020-vor

Orth, R., Denissen, J. M. C., Li, W., & Oh, S. (2023). Increasing water limitation of global ecosystems in a changing climate. *Copernicus Meetings*. https://doi.org/10.5194/egusphere-egu23-1422

Papagiannopoulou, C., Miralles, D. G., Decubber, S., Demuzere, M., Verhoest, N. E. C., Dorigo, W. A., & Waegeman, W. (2017). A non-linear Granger-causality framework to investigate climate–vegetation dynamics. *Geoscientific Model Development*, *10*(5), 1945–1960. https://doi.org/10.5194/gmd-10-1945-2017

Pelletier, J. D., Broxton, P. D., Hazenberg, P., Zeng, X., Troch, P. A., Niu, G. Y., et al. (2016). A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling. *Journal of Advances in Modeling Earth Systems*, *8*(1), 41–65. https://doi.org/10.1002/2015MS000526

Pinnington, E., Amezcua, J., Cooper, E., Dadson, S., Ellis, R., Peng, J., et al. (2021). Improving soil moisture prediction of a high-resolution land surface model by parameterising pedotransfer functions through assimilation of SMAP satellite data. *Hydrology and Earth System Sciences*, *25*(3), 1617–1641. https://doi.org/10.5194/hess-25-1617-2021

Pinnington, E., Quaife, T., Lawless, A., Williams, K., Arkebauer, T., & Scoby, D. (2020). The land variational ensemble data assimilation framework: LAVENDAR v1.0.0. *Geoscientific Model Development*, *13*(1), 55–69. https://doi.org/10.5194/gmd-13-55-2020

Pokhrel, Y., Hanasaki, N., Koirala, S., Cho, J., Yeh, P. J.-F., Kim, H., et al. (2012). Incorporating anthropogenic water regulation modules into a land surface model. *Journal of Hydrometeorology*, *13*(1), 255–269. https://doi.org/10.1175/JHM-D-11-013.1

Pokhrel, Y. N., Hanasaki, N., Wada, Y., & Kim, H. (2016). Recent progresses in incorporating human land–water management into global land surface models toward their integration into Earth system models. *WIREs Water*, *3*(4), 548–574. https://doi.org/10.1002/wat2.1150

Preimesberger, W., Scanlon, T., Su, C.-H., Gruber, A., & Dorigo, W. (2021). Homogenization of structural breaks in the global ESA CCI soil moisture multisatellite climate data record. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(4), 2845–2862. https://doi.org/10.1109/TGRS.2020.3012896

Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A. C., Müller, C., Arneth, A., et al. (2014). Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proceedings of the National Academy of Sciences*, *111*(9), 3268–3273. https://doi.org/10.1073/pnas.1222463110

Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., Boote, K. J., Thorburn, P., et al. (2013). The agricultural model intercomparison and improvement project (AgMIP): Protocols and pilot studies. *Agricultural and Forest Meteorology*, *170*, 166–182. https://doi.org/10.1016/j.agrformet.2012.09.011

Ryel, R. J., Caldwell, M. M., Yoder, C. K., Or, D., & Leffler, A. J. (2002). Hydraulic redistribution in a stand of Artemisia tridentata: Evaluation of benefits to transpiration assessed with a simulation model. *Oecologia*, *130*(2), 173–184. https://doi.org/10.1007/s004420100794

Sabot, M. E. B., De Kauwe, M. G., Pitman, A. J., Medlyn, B. E., Ellsworth, D. S., Martin-StPaul, N. K., et al. (2022). One stomatal model to rule them all? Toward improved representation of carbon and water exchange in global models. *Journal of Advances in Modeling Earth Systems*, *14*(4), e2021MS002761. https://doi.org/10.1029/2021MS002761

Sacks, W. J., Cook, B. I., Buenning, N., Levis, S., & Helkowski, J. H. (2009). Effects of global irrigation on the near-surface climate. *Climate Dynamics*, *33*(2–3), 159–175. https://doi.org/10.1007/s00382-008-0445-z

Samaniego, L., Thober, S., Kumar, R., Wanders, N., Rakovec, O., Pan, M., et al. (2018). Anthropogenic warming exacerbates European soil moisture droughts. *Nature Climate Change*, *8*(5), 421–426. https://doi.org/10.1038/s41558-018-0138-5

Seo, E., & Dirmeyer, P. A. (2022). Improving the ESA CCI daily soil moisture time series with physically based land surface model datasets using a Fourier time-filtering method. *Journal of Hydrometeorology*, *23*, 473–489. https://doi.org/10.1175/JHM-D-21-0120.1

Shah, H. L., Zhou, T., Huang, M., & Mishra, V. (2019). Strong influence of irrigation on water budget and land surface temperature in Indian subcontinental river basins. *Journal of Geophysical Research: Atmospheres*, *124*(3), 1449–1462. https://doi.org/10.1029/2018JD029132

Soltani, S. S., Fahs, M., Bitar, A. A., & Ataie-Ashtiani, B. (2022). Improvement of soil moisture and groundwater level estimations using a scale-consistent river parameterization for the coupled ParFlow-CLM hydrological model: A case study of the Upper Rhine basin. *Journal of Hydrology*, *610*, 127991. https://doi.org/10.1016/j.jhydrol.2022.127991

Strebel, L., Bogena, H. R., Vereecken, H., & Hendricks Franssen, H.-J. (2022). Coupling the community land model version 5.0 to the parallel data assimilation framework PDAF: Description and applications. *Geoscientific Model Development*, *15*(2), 395–411. https://doi.org/10.5194/gmd-15-395-2022

Sulis, M., Couvreur, V., Keune, J., Cai, G., Trebs, I., Junk, J., et al. (2019). Incorporating a root water uptake model based on the hydraulic architecture approach in terrestrial systems simulations. *Agricultural and Forest Meteorology*, *269–270*, 28–45. https://doi.org/10.1016/j.agrformet.2019.01.034

Sulis, M., Langensiepen, M., Shrestha, P., Schickling, A., Simmer, C., & Kollet, S. J. (2015). Evaluating the influence of plant-specific physiological parameterizations on the partitioning of land surface energy fluxes. *Journal of Hydrometeorology*, *16*(2), 517–533. https://doi.org/10.1175/JHM-D-14-0153.1

Tai, A. P. K., Martin, M. V., & Heald, C. L. (2014). Threat to future global food security from climate change and ozone air pollution. *Nature Climate Change*, *4*(9), 817–821. https://doi.org/10.1038/nclimate2317

Tang, J., Riley, W. J., & Niu, J. (2015). Incorporating root hydraulic redistribution in CLM4.5: Effects on predicted site and global evapotranspiration, soil moisture, and water storage. *Journal of Advances in Modeling Earth Systems*, *7*(4), 1828–1848. https://doi.org/10.1002/2015MS000484

Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., van Beek, R., Wada, Y., et al. (2013). Ground water and climate change. *Nature Climate Change*, *3*(4), 322–329. https://doi.org/10.1038/nclimate1744

Tian, W., Li, X., Cheng, G.-D., Wang, X.-S., & Hu, B. X. (2012). Coupling a groundwater model with a land surface model to improve water and energy cycle simulation. *Hydrology and Earth System Sciences*, *16*(12), 4707–4723. https://doi.org/10.5194/hess-16-4707-2012

Trugman, A. T., Medvigy, D., Mankin, J. S., & Anderegg, W. R. L. (2018). Soil moisture stress as a major driver of carbon cycle uncertainty. *Geophysical Research Letters*, *45*(13), 6495–6503. https://doi.org/10.1029/2018GL078131

Tumbo, S. D., Mutabazi, K. D., Mourice, S. K., Msongaleli, B. M., Wambura, F. J., Mzirai, O. B., et al. (2020). Integrated assessment of climate change impacts and adaptation in agriculture: The case study of the Wami river sub-basin, Tanzania. In J. I. Matondo, B. F. Alemaw, & W. J. P. Sandwidi (Eds.), *Climate variability and change in Africa: Perspectives, experiences and sustainability* (pp. 115–136). Springer International Publishing. https://doi.org/10.1007/978-3-030-31543-6_10

Ukkola, A. M., Kauwe, M. G. D., Pitman, A. J., Best, M. J., Abramowitz, G., Haverd, V., et al. (2016). Land surface models systematically overestimate the intensity, duration and magnitude of seasonal-scale evaporative droughts. *Environmental Research Letters*, *11*(10), 104012. https://doi.org/10.1088/1748-9326/11/10/104012

Urban, D., Roberts, M. J., Schlenker, W., & Lobell, D. B. (2012). Projected temperature changes indicate significant increase in interannual variability of U.S. maize yields. *Climate Change*, *112*(2), 525–533. https://doi.org/10.1007/s10584-012-0428-2

Van der Velde, R., Colliander, A., Pezij, M., Benninga, H.-J. F., Bindlish, R., Chan, S. K., et al. (2021). Validation of SMAP L2 passive-only soil moisture products using upscaled in situ measurements collected in Twenty, The Netherlands. *Hydrology and Earth System Sciences*, *25*(1), 473–495. https://doi.org/10.5194/hess-25-473-2021

Victoria Government Data Directory. (2018). Agriculture Victoria research division in the department of economic development, jobs, transport, and resources, spatial Sciences group: Victorian land use information system 2016. [Dataset]. https://doi.org/10.4226/92/590abbe6ea3f1

Viovy, N. (2018). CRUNCEP version 7 - atmospheric forcing data for the community. *Land Model*. https://doi.org/10.5065/PZ8F-F017

Wada, Y., van Beek, L. P. H., & Bierkens, M. F. P. (2012). Nonsustainable groundwater sustaining irrigation: A global assessment. *Water Resources Research*, *48*(6). https://doi.org/10.1029/2011WR010562

Wada, Y., van Beek, L. P. H., van Kempen, C. M., Reckman, J. W. T. M., Vasak, S., & Bierkens, M. F. P. (2010). Global depletion of groundwater resources. *Geophysical Research Letters*, *37*(20). https://doi.org/10.1029/2010GL044571

Wada, Y., Wisser, D., Eisner, S., Flörke, M., Gerten, D., Haddeland, I., et al. (2013). Multimodel projections and uncertainties of irrigation water demand under climate change. *Geophysical Research Letters*, *40*(17), 4626–4632. https://doi.org/10.1002/grl.50686

Weihermüller, L., Lehmann, P., Herbst, M., Rahmati, M., Verhoef, A., Or, D., et al. (2021). Choice of pedotransfer functions matters when simulating soil water balance fluxes. *Journal of Advances in Modeling Earth Systems*, *13*(3), e2020MS002404. https://doi.org/10.1029/2020MS002404

White, J. W., Hunt, L. A., Boote, K. J., Jones, J. W., Koo, J., Kim, S., et al. (2013). Integrated description of agricultural field experiments and production: The ICASA Version 2.0 data standards. *Computers and Electronics in Agriculture*, *96*, 1–12. https://doi.org/10.1016/j.compag.2013.04.003

Xia, Q., Liu, P., Fan, Y., Cheng, L., An, R., Xie, K., & Zhou, L. (2022). Representing irrigation processes in the land surface-hydrological model and a case study in the Yangtze river basin, China. *Journal of Advances in Modeling Earth Systems*, *14*(7), e2021MS002653. https://doi.org/10.1029/2021MS002653

Yan, B., & Dickinson, R. E. (2014). Modeling hydraulic redistribution and ecosystem response to droughts over the Amazon basin using Community Land Model 4.0 (CLM4). *Journal of Geophysical Research: Biogeosciences*, *119*(11), 2130–2143. https://doi.org/10.1002/2014JG002694

Yassin, F., Razavi, S., Elshamy, M., Davison, B., Sapriza-Azuri, G., & Wheater, H. (2019). Representation and improved parameterization of reservoir operation in hydrological and land-surface models. *Hydrology and Earth System Sciences*, *23*(9), 3735–3764. https://doi.org/10.5194/hess-23-3735-2019

Yuan, X., Xie, Z., Zheng, J., Tian, X., & Yang, Z. (2008). Effects of water table dynamics on regional climate: A case study over EasSt Asian monsoon area. *Journal of Geophysical Research*, *113*(D21). https://doi.org/10.1029/2008JD010180

Zhao, H., Montzka, C., Baatz, R., Vereecken, H., & Franssen, H.-J. H. (2021). The importance of subsurface processes in land surface modeling over a temperate region: An analysis with SMAP, cosmic ray neutron sensing and triple collocation analysis. *Remote Sensing*, *13*(16), 3068. https://doi.org/10.3390/rs13163068

Zheng, Z., & Wang, G. (2007). Modeling the dynamic root water uptake and its hydrological impact at the Reserva Jaru site in Amazonia. *Journal of Geophysical Research*, *112*(G4). https://doi.org/10.1029/2007JG000413