Field Crops Research 332 (2025) 110032

Contents lists available at ScienceDirect

Field Crops Research

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

Wheat crop models underestimate drought stress in semi-arid and Mediterranean environments

H. Webber [a,b,*], D. Cooke [a], C. Wang [a], S. Asseng [c], P. Martre [d], F. Ewert [a,e], B. Kimball [f], G. Hoogenboom [g], S. Evett [h], A. Chanzy [i], S. Garrigues [j], A. Olioso [i,k], K.S. Copeland [l], J.L. Steiner [m], D. Cammarano [n], Y. Chen [o,p], M. Crépeau [q], E. Diamantopoulos [r], R. Ferrise [s], L. Manceau [d], T. Gaiser [e], Y. Gao [g,t], S. Gayler [u], J.R. Guarin [v], T. Hunt [w], G. Jégo [q], G. Padovan [s], E. Pattey [x], D. Ripoche [y], A. Rodríguez [z], M. Ruiz-Ramos [aa], V. Shelia [g], A.K. Srivastava [a,e], I. Supit [ab], F. Tao [o,p], K. Thorp [f], M. Viswanathan [ac], T. Weber [u,ad], J. White [g]

[a] Leibniz Centre for Agricultural Landscape Research (ZALF), Müncheberg, Germany
[b] Brandenburg University of Technology, Cottbus, Germany
[c] Digital Agriculture, HEF World Agricultural Systems Center, Technical University of Munich, Freising, Germany
[d] LEPSE, Univ Montpellier, INRAE, Institut Agro Montpellier, Montpellier, France
[e] Crop Science, Institute of Crop Science and Resource Conservation, University of Bonn, Bonn, Germany
[f] US Arid-Land Agricultural Research Center, USDA-ARS, Maricopa, USA
[g] Agricultural and Biological Engineering Dept., University of Florida, Gainesville, USA
[h] USDA ARS, Conservation and Production Research Laboratory, USDA-ARS, Bushland, USA
[i] UMR EMMAH, INRAE, Avignon Université, Avignon, France
[j] European Centre for Medium Range Weather Forecasts, Shinfield Park, Reading, United Kingdom
[k] Unité de Recherche écologie des Forêts Méditerranéennes (URFM), INRAE, Avignon, France
[l] USDA ARS Conservation & Production Research Laboratory, Bushland, TX, USA
[m] Agronomy Department, Kansas State University, Manhattan, KS, USA
[n] Department of Agroecology, iClimate, CBIO, Aarhus University, Tjele, Denmark
[o] Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China
[p] College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, China
[q] Science and Technology Branch, Agriculture and Agri-Food Canada, Quebec City, Quebec, Canada
[r] University of Bayreuth, Bayreuth, Germany
[s] Department of Agri-food Production and Environmental Sciences (DISPAA), University of Florence, Florence, Italy
[t] Institute for Sustainable Food Systems, University of Florida, Gainesville, USA
[u] Institute of Soil Science and Land Evaluation, Hohenheim University, Stuttgart, Germany
[v] NASA Goddard Institute for Space Studies, New York, USA
[w] University of Guelph, Guelph, Canada
[x] Central Experimental Farm, Agriculture and Agri-Food Canada, Ottawa, Canada
[y] AgroClim, INRAE Domaine Saint-Paul, Avignon, France
[z] Department of Economic Analysis and Finances, University of Castilla-La Mancha, Toledo, Spain
[aa] CEIGRAM, Technical University of Madrid, Madrid, Spain
[ab] Earth Systems and Global Change Group, Wageningen University, Wageningen, the Netherlands
[ac] Julius Kühn-Institute (JKI), Institute for Strategies and Technology Assessment, Germany
[ad] Soil Science Section, Faculty of Organic Agricultural Sciences, University of Kassel, Witzenhausen, Germany

ARTICLE INFO

Keywords:
Crop models
Evapotranspiration
Wheat
Drought stress
Climate risk

ABSTRACT

Under climate change and increasingly extreme weather, projections of water demand and drought stress from process-based crop models can inform risk management and adaptation strategies. Previous studies investigating maize crop models demonstrated considerable error in the simulation of water use, and no similar evaluation of wheat crop models exists. The aims of this study were to (1) evaluate wheat crop models' performance in reproducing observed daily evapotranspiration (ET) for Mediterranean and semi-arid environments, and (2) identify factors and processes associated with model error and uncertainty. These were assessed with an

* Corresponding author at: Leibniz Centre for Agricultural Landscape Research (ZALF), Müncheberg, Germany.
E-mail address: webber@zalf.de (H. Webber).

https://doi.org/10.1016/j.fcr.2025.110032
Received 7 March 2025; Received in revised form 7 May 2025; Accepted 7 June 2025
Available online 13 June 2025
0378-4290/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

ensemble of wheat crop models for two experiments, one conducted in Bushland, Texas, USA (three seasons, deficit and full irrigation) and another in Avignon, France (four rainfed seasons) with winter bread and durum wheat, respectively. Models were calibrated with all observed data for crop growth. The model ensemble median underestimated water use in all environments evaluated, suggesting a systematic bias. The relative error in underestimating daily ET was constant across levels of atmospheric evaporative demand; therefore, the absolute error was greater for days with larger evaporative demand. This implies errors in the soil water balance increase more rapidly under high evaporative demand conditions. Using a potential versus reference crop evapotranspiration approach did not explain relative model performance. However, the sensitivity analysis indicated that simulation of atmospheric evaporative demand terms explained much more uncertainty in seasonal water use than terms related to soil depth or root growth. Errors in simulated leaf area index were associated with errors in daily simulated ET, but the relationship varied with the growth stage. Collectively, the results suggest the need to improve simulation of atmospheric ET demand to avoid underestimating projected impacts of drought or required water resource availability for viable production systems.

1. Introduction

Drought and the resulting crop water stress are major constraints to agricultural production (Lesk et al., 2016). Water scarcity limits crop yields in many regions (Van Ittersum and Rabbinge, 1997), and drought can result in widespread yield failures even in regions normally characterized by good rainfall conditions (Webber et al., 2020). Drought is projected to increase under climate change in many regions (IPCC, 2022) with consequent negative implications for food security (Wheeler and Von Braun, 2013), water availability in natural ecosystems (Zhao et al., 2022), water and infrastructure requirements in irrigated systems (Evett et al., 2020) and carbon and nitrogen cycles in managed land (Liu et al., 2020; Stocker et al., 2019).

Process-based crop growth models have been widely applied to project climate change impacts on crop yields (Rezaei et al., 2023; White et al., 2011) and as components of integrated modelling approaches to inform wider system adaptation and mitigation strategies (Peng et al., 2020). However, there is a high degree of uncertainty in model projections of crop yield (Asseng et al., 2013) and water use (Cammarano et al., 2016). In response, considerable effort has been invested in improving model responses to warmer temperatures (Maiorano et al., 2017; Wang et al., 2017), heat stress (Gabaldón-Leal et al., 2016) and the combination of heat and drought stress (Webber et al., 2017, 2018). Kimball et al. (2019) demonstrated large errors and variation in the simulation of crop water use and drought stress for maize crop models. However, very few similar studies have been undertaken to evaluate the performance of wheat crop models in simulating water use. One of the few such examples contrasted model skill in simulating wheat crops in West Texas, USA, and Hunan Province, PR China (Kang et al., 2009). In that study, CROPWAT, MODWht, and CERES-Wheat were compared and found to overestimate ET at low water use rates and underestimate ET at large water use rates. Poor estimation of leaf area index was found to be an important factor in explaining poor model performance, and model revisions were recommended. Wheat crop model skill in simulation of daily evapotranspiration (ET) has implications for capturing daily responses to water deficits and the cumulative effects on soil water supply and subsequent availability (Webber et al., 2022). Additionally, as transpiration is a key determinant of canopy temperature, correct simulation of daily ET is also expected to be central to robust projections of crop temperature responses.

Improving simulation of crop water use is challenged by a lack of observational data across environments. Comprehensive ET and soil moisture datasets for a particular site and cultivar across multiple growing seasons are limited compared to the availability of similar datasets focusing on crop yield and aboveground biomass. Further, it is difficult to prioritize what to improve as multiple processes are involved in crop water use. Solar radiation and atmospheric conditions such as temperature and humidity determine the evaporative demand and resulting water use when soil water availability and plant hydraulics do not limit soil evaporation (E) nor transpiration (T), respectively. The partitioning between E and T is largely dependent on leaf area index. Evaporation is influenced by soil conditions, surface residues, and water availability. In contrast, T is influenced by plant and leaf structural characteristics, as well as root water extraction and stomatal behavior, which in turn are affected by soil water availability. Exposure to ozone, CO_2, pests and disease can also affect T. In turn, soil water extraction will be affected by root growth patterns and age, as well as soil properties and crop response to water deficit. Each of these processes, factors and their interactions can be simulated in various ways, and with different degrees of complexity. The relative importance of these individual processes will vary across environments and seasons as well as with the crop genotype and the development stage of the crop.

One particularly important distinction in the approach to simulate daily ET is whether crop models use reference crop ET estimates (ET_0) or rely on direct estimates of potential ET (ET_P). The latter approach allows for the coupling of T and atmospheric conditions but must be locally calibrated, which is rarely performed or possible (Allen et al., 1998). Additionally, while E and T are distinct processes, they are difficult to determine individually at canopy scale and are usually reported as a combined ET value. Finally, crop models simulate daily ET in interaction

Table 1
Site and growing season-specific data for bushland (TX) and avignon (FR). treatment naming follows the convention: sowing year-water status-site, where water status refers to either full (full) or deficit (dfct) irrigation or rainfed conditions. Management data displayed include cultivar grown and total irrigation amount applied (Irr.). the weather data are summarized for the growing season duration as maximum (Max. T) and minimum (Min. T) air temperatures, and rain. The growing season duration is in days. Biomass and growth data includes maximum leaf area index (LAI), maximum above-ground biomass (AGB) and final grain yield (Yield).

Treatment name	Cultivar	Growing season (d)	Rain (mm)	Irr. (mm)	Max. T (°C)	Min. T (°C)	Max. LAI	AGB (tha^{-1})	Yield (t ha^{-1})
1989-full-TX	TAM200	257	181	532	41.5	−21.6	3.7	16.3	4.5
1989-dfct- TX	TAM200	249	190	122	41.5	−21.6	2.4	8.2	1.9
1991-full- TX	TAM107	277	554	404	35.5	−15	7.1	22.0	5.8
1991-dfct- TX	TAM107	264	448	98	35.5	−15	5.9	13.9	3.2
1992-full- TX	MESA	268	287	477	35.4	−15.4	4.1	18.3	6.4
1992-dfct- TX	MESA	273	296	253	35.4	−15.4	4.0	15.1	5.0
2003-rain-FR	Durum Artimon	205	335	0	33.7	−3.5	2.7	17.7	7.1
2005-rain-FR	Durum Acalou	198	169	20	30.9	−4.9	5.5	15.3	6.1
2009-rain-FR	Durum Dakter	183	372	0	29.2	−6.9	6.4	16.7	6.1
2011-rain-FR	Durum Isildur	233	394	0	34.0	−7.2	NA	10.4	4.2

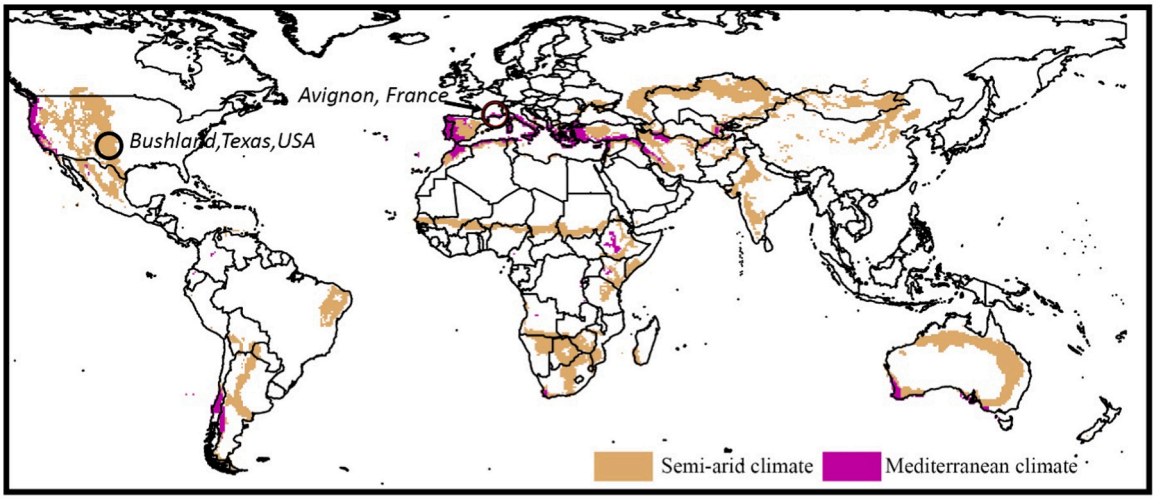

Fig. 1. Location of the two experimental sites indicating (a) the global distribution of semi-arid and Mediterranean climate zones. Note, Antarctica and the northernmost parts of the artic have been removed from the figure to optimize the layout as these regions have no wheat production.

with many other processes related to growth (e.g., leaf area index, phenology, biomass and root growth, nutrient use, and other possible crop stress), so errors in any of these processes may propagate to errors in simulated water use.

Against this background, the overarching aim of this study was to support model improvement by identifying critical processes for semi-arid and Mediterranean environments that contribute to model error in the simulation of crop water use. We focus on semi-arid and Mediterranean environments as drought stress and the need for irrigation is common in these regions, and their extent is projected to increase with climate change (Huang et al., 2017; Koppa et al., 2024). The specific objectives are to (1) evaluate model performance and ensemble variation in reproducing observed daily ET of wheat; and (2) identify factors and processes associated with model error and uncertainty in simulation of daily ET across environments. These objectives were assessed with an ensemble of wheat crop models for two experiments, one conducted in Bushland, Texas, USA and the other in Avignon, France, with winter bread (*Triticum* aestivum L.) and durum wheat (*T. durum* Desf.), respectively.

2. Materials and methods

2.1. Study regions

The study was conducted using datasets from Bushland, Texas, USA, and Avignon, France (Table 1, Fig. 1). The following sections describe the individual datasets. Previous publications by Evett et al. (1994), Evett et al. (2016), Howell et al. (1993), Howell et al. (1995b) and Howell et al. (2006) provide more detailed descriptions of the experimental sites and observational data collection procedures.

2.1.1. Bushland, texas

Data from Bushland, Texas (35°11'N latitude; 102°06'W longitude; 1170 m elevation) included weather, soil properties, crop management, phenological and biomass measurements and other observational data from three growing seasons at a USDA-ARS site (planting years 1989, 1991 and 1992). Each year, two fields were planted with winter wheat: one managed under full irrigation to avoid water stress, the other under partial irrigation for a total of six treatments. Full irrigation plots were managed to maintain 80 % or greater extractable water in the soil profile, while plots under partial irrigation were irrigated only at the beginning of the growing season to ensure crop establishment. Different winter bread wheat cultivars were planted each year but remained the same between treatments for a given year. Daily ET was measured using weighing lysimeters. Soil water content was determined using six in-field neutron tubes, two tubes in the lysimeter (to depth of 190 cm) and four across the field (to 230 cm). For more information on the original experiments, see Evett et al. (1994), Evett et al. (2022a), Howell et al. (1993), Howell et al. (1995b) and Howell et al. (2006). The data are available at https://doi.org/10.15482/USDA.ADC/1527912. Quality control of weather data is described in Evett et al. (2018). Quality control of lysimeter ET data is described in Evett et al. (2012), Evett et al. (2024), Howell et al. (1995a) and Marek et al. (2014). Quality control of water content data is described in Evett et al. (2003), Evett et al. (2008), Evett et al. (2022b) and Evett and Steiner (1995).

2.1.2. avignon, france

Data from Avignon, France were collected during a 12-year crop rotation at an INRAE research site (43°55'00.4" N latitude; 4°52'41.0" E longitude; 32 m altitude). Rotations were split into summer and winter crops, with fallow periods in between. A subset of the full dataset pertaining to winter durum wheat growing seasons was provided (planting years 2003, 2005, 2009, and 2011). There were no distinct treatments between growing seasons, though some years differed in terms of cultivar planted or previous crop. Four cultivars were used throughout the experiment (Table 1). Crops were primarily rainfed, although a small amount of irrigation was administered in one year to ensure crop establishment. The groundwater depth was at depth of greater than 4 m and groundwater contributions to the rootzone are considered negligible. Daily ET rates were estimated from an eddy covariance system, which had a northward orientation into the main wind direction (Garrigues et al., 2015). Gas flux measurements were recorded twice per hour and summed to compute cumulative ET. Soil moisture was measured using neutron probes up to 150–190 cm depth, depending on the year. Based on results in Garrigues et al. (2015) in which relatively large variation in field capacity values determined with gravimetric samples as compared to those determined with pressure plates in the lab, for the layer 0.4–0.8 m, values were modified from.36 to 0.32; and for layer 0.8–1.2 m, value was modified from 0.35 to 0.3 in much better agreement with the observed soil moisture values. More information on the full 12-year crop rotation experiment and soil properties are found in Garrigues et al. (2015) and Bruckler et al. (2004), respectively.

2.2. Simulation protocol

Participation in the model comparison was open to all modellers in the Agricultural Model Intercomparison and Improvement Project (AgMIP) Wheat group (Rosenzweig et al., 2013). To participate,

Table 2
Description of the factors and processes considered in the sensitivity analyses, the range of variation explored across 5-equally spaced levels and the number of models able to vary the factor or process.

Daily reference crop ET (ET_0)		Daily potential ET (ET_P)		Daily soil water evaporation (E)		Daily root depth increment (RDI)		Max rooting depth (RDM)	
Range	No. of models	Range	No. of models	Range	No. of models	Range	No. of models	Range	No. of models
−20 %-+ 20 %	9	−20 %-+ 20 %	17	−20 %-+ 20 %	17	−10 %-+ 10 %	17	−20 %−0 %	17

modellers were required to be able to modify the source code of their model to vary the simulated values of daily rates of ET_0 and/or ET_P, soil water evaporation demand, root depth increment and the maximum rooting depth. For results presented in this study, modellers had access to all calibration data. The protocols, the reporting templates and all data available to modellers at published at: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/HDKKAL.

The data was provided to modellers in the form of spreadsheets following ICASA data standards (White et al., 2013). Datasets for each site included detailed field observations, including dates of key phenological stages (e.g., emergence, anthesis, maturity and/or harvest date), time series of leaf area index (LAI) and above-ground biomass, and end-of-season yields and above-ground biomass, as well as information to characterize the site such as soil texture by horizon, crop management, and weather data. For both sites, soil moisture observations were available approximately every two to three weeks. The availability of data on daily crop evapotranspiration rates differed between the sites. At Bushland, data was available daily from the weighing lysimeters. At Avignon, observations of daily ET were only available on between 3 % and 10 % of days in the growing season due to a rigorous quality control protocol (Mauder et al., 2013). We considered only observations for the 24-hour period during which each individual 30-minute measurement was valid.

Modellers were given a detailed protocol to help them understand the main aim and study design, as well as to provide detailed instructions on how to conduct the simulations. Models were to have been calibrated to within + /- 5 days of the observed data, even if it required year-specific phenology calibration where varieties coincided. Simulations for each treatment were to be started on the date when initial conditions for soil water content were provided, usually a few days before seeding. Simulations were to be terminated when the models simulated physiological maturity. Modellers were provided with templates to report on a series of daily and cumulative variables. Cumulative variables were to be calculated from the day of initial conditions to simulated maturity. Modellers were instructed to calibrate their model using all data provided and to refrain from consulting publications from the experiments. Additionally, any parameters that were adjusted in the calibration of the models regarding soil water dynamics or daily ET were reported.

2.2.1. Sensitivity analyses

A sensitivity analysis was performed in each of the low and full data phases to explore which underlying processes or factors had the largest influence on crop water use and the degree to which this influence varied between models. Following the method described by Webber et al. (2016), processes and factors, further referred to as components, relevant to crop water use that could be simulated and modified by the participating modellers were identified. The selected components were ET_0, ET_P, E, daily root depth increment (RDI) and max rooting depth (RDM). Each of these components was varied to provide five distinct levels between the ranges given in Table 2. Simulations were conducted for all factorial combinations. The required changes were implemented by individual modellers through changes in the source code (or potentially a parameter file or input variable for RDI and RDM, respectively). Modellers were requested to implement the changes for each component at the line of source code where the final variable (e.g., ET_0, ET_p, or E) was calculated by using an appropriate multiplicative factor corresponding to those in Table 2 (e.g., 0.80, 0.90, 1.00, 1.10 and 1.20 for ET0, ETP, E, RDM and 0.90, 0.95, 1.0, 1.05 and 1.10). This approach avoided the need to harmonize methods or model algorithms for this explorative stage. Rather, the simulated values were directly modified by multiplying the daily estimate of the respective component by the multiplicative factors stated above. Given that some models do not estimate ET_0 and rather directly estimate ET_P, two sensitivity analyses were performed (one with and one without ET_0 models). Details on analytical methods are provided in Section 2.3.3.

2.2.2. Participating models

Thirteen modelling groups participated in the study (Table 3). Many of the groups shared a common base model or model framework. For example, DSSAT model variants were used by four groups, whereas variants of the SIMPLACE<Lintul5 > , Expert-N and STICS models were used by two groups each. Each of the four DSSAT-based model variants performed a special set of simulations. They each implemented two approaches to simulate daily crop evapotranspiration (an ET_0 and ET_P approach). The two approaches for simulating ET_0 and ET_P were the FAO-56 Penman-Monteith combination equation for a grass reference crop (Allen et al., 1998) and the Priestley-Taylor ET_P approach (Priestley and Taylor, 1972), respectively. While the DSSAT groups also considered two approaches to simulate soil water evaporation, the current study used only their simulations with the two-stage Ritchie approach (Ritchie, 1972).

Notably, four modeling groups used DSSAT-based models have participated in this research. They contributed a total of four distinct DSSAT model structures: DSSAT CSM-CERES-Wheat, DSSAT-CROPSIM, and DSSAT-N-Wheat, with the DSSAT CSM-CERES-Wheat model calibrated independently by two different teams. This means that while only three versions of DSSAT models were used, they contributed four model entries due to the independent calibration strategies applied to the CSM-CERES-Wheat model. Since each DSSAT model was simulated with two ET approaches, the total number of DSSAT simulations was eight. When combined with simulations from the other nine models, this got 17 model simulations in total.

2.3. Analysis methods

2.3.1. Developmental stages

Three developmental stages relevant for crop water use were defined to examine the models' performance in simulating ET for different periods of relevance for partitioning between T and E. The stages were determined based on crop LAI. The three stages were (a) early vegetative (pre-anthesis and LAI < 0.5); (b) full cover ((LAI ≥ 0.5 and maximum LAI not yet reached) or (maximum LAI reached and LAI $\geq$ maximum LAI - 1)); and (c) senescence (after maximum LAI reached and LAI $\leq$ maximum LAI – 1). As observed LAI was only available for a few dates in each experiment, for each treatment, the median of simulated LAI was used as a proxy (SI Figs. S2 and S3) together with information about the maximum LAI and whether anthesis had already been achieved to estimate a simplified growth stage (early vegetative, full cover, and senescence) for each day. For individual models, the model-specific growth stage was determined considering the model simulated LAI, maximum LAI, and date of maximum LAI. A model was deemed correct in simulating phenology when the model median and model-specific estimates of growth stages were the same.

Table 3
Overview of the participating models and a description of the relevant modelling approaches for simulating crop water use, particularly the components varied in the sensitivity analysis (I.e., reference crop evapotranspiration, potential evapotranspiration, soil water evaporation, and root growth).

Model name (key references)	Model acronym	ET0 method	Potential ET	Partitioning between T and E	Potential soil water evaporation	Root water uptake method	Water stress effects
DSSAT CSM-CERES-Wheat (Jones et al., 2003**)**	DIFR	FAO56 Penman-Monteith combination equation for a grass reference crop (Allen et al., 1998)	FAO56 Crop coefficient approach (Allen et al., 1998)	function of crop LAI	Two-stage Ritchie model (Ritchie et al., 2009)	Root water uptake is computed by soil layer (cm/d), in a water balance from with daily weather values, all soil properties, current soil water content, LAI and root length density for each layer (Ritchie, 1998).	Stomatal conductance reduced. Photosynthesis decreased proportionally to transpiration, reduced plant turgor and expansive growth. Senescence and biomass partitioning are altered
DSSAT CSM-CERES-Wheat(Jones et al., 2003**)**	D1RR	NA	Priestley-Taylor potential ET	function of crop LAI	Two-stage Ritchie model (Ritchie et al., 2009)	Root water uptake is computed by soil layer (cm/d), in a water balance from with daily weather values, all soil properties, current soil water content, LAI and root length density for each layer (Ritchie, 1998).	Stomatal conductance reduced. Photosynthesis decreased proportionally to transpiration, reduced plant turgor and expansive growth. Senescence and biomass partitioning are altered
DSSAT CSM-CERES-Wheat (Jones et al., 2003**)**	D3FR	FAO56 Penman-Monteith combination equation for a grass reference crop (Allen et al., 1998)	FAO56 Crop coefficient approach (Allen et al., 1998)	function of crop LAI	Two-stage Ritchie model (Ritchie et al., 2009)	Root water uptake is computed by soil layer (cm/d), in a water balance from with daily weather values, all soil properties, current soil water content, LAI and root length density for each layer (Ritchie, 1998).	Stomatal conductance reduced. Photosynthesis decreased proportionally to transpiration, reduced plant turgor and expansive growth. Senescence and biomass partitioning are altered
DSSAT CSM-CERES-Wheat(Jones et al., 2003**)**	D3RR	NA	Priestley-Taylor potential ET	function of crop LAI	Two-stage Ritchie model (Ritchie et al., 2009)	Root water uptake is computed by soil layer (cm/d), in a water balance from with daily weather values, all soil properties, current soil water content, LAI and root length density for each layer (Ritchie, 1998).	Stomatal conductance reduced. Photosynthesis decreased proportionally to transpiration, reduced plant turgor and expansive growth. Senescence and biomass partitioning are altered.
DSSAT-CROPSIM (Jones et al., 2003**)**	DRFR	FAO56 Penman-Monteith combination equation for a grass reference crop (Allen et al., 1998)	FAO56 Crop coefficient approach (Allen et al., 1998)	T function of crop LAI	Two-stage Ritchie model (Ritchie et al., 2009)	Root water uptake is computed by soil layer (cm/d), in a water balance from with daily weather values, all soil properties, current soil water content, LAI and root length density for each layer (Ritchie, 1998).	Stomatal conductance reduced. Photosynthesis decreased proportionally to transpiration, reduced plant turgor and expansive growth. Senescence and biomass partitioning are altered
DSSAT-CROPSIM(Jones et al., 2003**)**	DRRR	NA	Priestley-Taylor potential ET	function of crop LAI	Two-stage Ritchie model (Ritchie et al., 2009)	Root water uptake is computed by soil layer (cm/d), in a water balance from with daily weather values, all soil properties, current soil water content, LAI and root length density for each layer (Ritchie, 1998).	Stomatal conductance reduced. Photosynthesis decreased proportionally to transpiration, reduced plant turgor and expansive growth. Senescence and biomass

(*continued on next page*)

Table 3 (*continued*)

Model name (key references)	Model acronym	ET0 method	Potential ET	Partitioning between T and E	Potential soil water evaporation	Root water uptake method	Water stress effects
							partitioning are altered
DSSAT-N-Wheat	DNFR	FAO56 Penman-Monteith combination equation for a grass reference crop (Allen et al., 1998)	Minimum of Priestley-Taylor potential ET and transpiration efficiency of simulated biomass growth	Function of crop LAI	Two-stage Ritchie model (Ritchie et al., 2009)	Based on the simulated crop water demand, root length density and soil water content per soil layer (Asseng et al., 2004)	Soil water deficit factor (0−1) based on the root zone available water to decrease photosynthesis (daily carbohydrate production), root elongation, leaf expansion and tillering, and accelerate leaf senescence
DSSAT-N-Wheat	DNRR	NA	Minimum of Priestley-Taylor (Priestley and Taylor, 1972) with Ritchie soil water E (Ritchie et al., 2009) and transpiration efficiency of simulated biomass growth	Function of crop LAI	Two-stage Ritchie model (Ritchie et al., 2009)	Based on the simulated crop water demand, root length density and soil water content per soil layer (Asseng et al., 2004)	Soil water deficit factor (0−1) based on the root zone available water to decrease photosynthesis (daily carbohydrate production), root elongation, leaf expansion and tillering, and accelerate leaf senescence
Lintul4-vshs_c	L4XX	NA	Penman * CFEV	Function of the LAI and an extinction coefficient for global radiation (Goudriaan, 1977; Ritchie, 1972)	Based on Penman following the concepts of Goudriaan (1977) and Ritchie (1972)	Water absorption depends on the actual water content in the rootzone that can be absorbed. This is root depth dependent (Similar to Wofost: Supit et al.: 1994)	1/ Assimilation choking function depending on Max (waterstress, NitrogenStress). 2/ Similar to Wofost below development stage 0.75 Leaf extension is reduced
MCWLA-Wheat	MCXX	FAO56 Penman-Monteith combination equation for a grass reference crop (Allen et al., 1998)		function of crop LAI, crop demand and water supply (Tao et al., 2009)	Based on upper 20 cm soil layer and equilibrium ET (Ritchie et al., 1985)	Based on the available water in the root zone, root density and a parameter of extraction power (Haxeltine and Prentice, 1996)	Result in a reduced rate of increase in LAI and an increased rate of senescence (Tao and Zhang, 2013)
Expert-N-VGM (Gayler et al., 2013**;** Priesack et al., 2006**)**	NVXX	FAO Penman-Monteith method	ETpot=ET0 * Kc, Kc = f(DevStage)	function of crop LAI: Tpot= ETpot*f(LAI)	Epot= [ETpot*(1-f (LAI)]	Proportional to root length density, limited by availability of soil water, maximum root water uptake capacity, or potential transpiration, similar to (Jones et al., 1986)	If Tact<Tpot: Water stress effects on leaf expansion, photosynthesis, and assimilate partitioning (various response functions)
SIMPLACE<LINTUL−5, Hargreaves > (Enders et al., 2023**)**	L6XX	FAO56 Hargreaves (Allen et al., 1998)	FAO56 Dual crop coefficient approach (Allen et al., 1998)	FAO56 Dual Crop Coefficient approach, function of LAI (Allen et al., 1998)	Initially energy limited, then water limited in upper soil profile with FAO56 Dual Crop Coefficient approach (Allen et al., 1998)	Proportional to available soil water (which depends on root depth) and a maximum water extraction rate (Feddes, 1978)	Increases partitioning to roots, and reduces both radiation use efficiency and rate of exponential LAI growth
SIRIUSQUALITY	SQXX	NA	Resistive model of Penman (Jamieson et al., 1995)	T and E calculated separately following Ritchie (1972)	Two phase model of Ritchie (1972)	A proportion of available soil water from each soil layer can be extracted by the plant on any day. This proportion decreases with soil depth, mimicking the decrease of root length density with soil depth (Semenov et al., 2009)	Decreasing transpiration and radiation use efficiency, slowing leaf growth and accelerating leaf senescence

(*continued on next page*)

Table 3 (*continued*)

Model name (key references)	Model acronym	ET0 method	Potential ET	Partitioning between T and E	Potential soil water evaporation	Root water uptake method	Water stress effects
SSM-Wheat	SSXX	NA	Priestley-Taylor modified model (Priestley and Taylor, 1972; Ritchie, 1998)	The model calculates Potential ET and potential soil water evaporation. Potential crop transpiration is the difference between them	Two-stage Ritchie model	Daily transpiration rate is calculated from crop biomass, transpiration efficiency coefficient and vapor pressure deficit. Water uptake is proportional to the soil water content in the root zone, with less water being transpired from drier layers	Based on the root zone available water, drought and waterlogging stress factors affecting growth, leaf development and phenology are calculated
STICS-CC	STCC		Product of Penman method (Penman, 1948) and a crop coefficient-based (Brisson et al., 1992)	Function of the LAI	Two phases model of Brisson and Perrier (1991), based on concepts put forward by Ritchie (1972)	Water absorption in the root zone is distributed cm per cm, according to 2 factors: the effective root density profile and the available water content	Delaying emergence and crop development, decreasing transpiration and radiation use efficiency, slowing root growth in depth and leaf growth, accelerating leaf senescence
STICS-SW	STSW	NA	Resistive model of Shuttleworth and Wallace (Brisson et al., 1998)	Function of the LAI, the canopy geometry and the radiation quality (direct vs diffusive)	Two phase model of Brisson et al. (1998) based on concepts put forward by Ritchie (1972)	Water absorption in the root zone is distributed cm per cm, according to 2 factors: the effective root density profile and the available water content	Delaying emergence and crop development, decreasing transpiration and radiation use efficiency, slowing root growth in depth and leaf growth, accelerating leaf senescence
Expert-N-BW (Priesack et al., 2006**;** Weber et al., 2019**)**	NBXX	FAO56 Penman-Monteith combination equation for a grass reference crop (Allen et al., 1998)	Product of ET0 and crop coefficient, which varies with development stage	function of crop LAI	Function of crop LAI	Simialar as Expert-N-VGM but with a set of soil hydraulic property function specifically developed to physically comprehensively represent dry soil hydraulic properties (Weber et al., 2019, 2020)	If Tact<Tpot: Water stress effects on leaf expansion, photosynthesis, and assimilate partitioning (various response functions)

* A total of 13 unique models were included in this study. The DSSAT group contributed three model structures—DSSAT CSM-CERES-Wheat, DSSAT-CROPSIM, and DSSAT-N-Wheat. The DSSAT CSM-CERES-Wheat model was calibrated independently by two separate teams, and is therefore considered as two distinct contributions, resulting in four DSSAT model entries. Each of these four DSSAT models was simulated using two alternative approaches to estimate evapotranspiration (ET), leading to a total of eight DSSAT-based simulations. Combined with the nine simulations from the other modeling groups, this results in 17 model simulations in total

**NA if not applicable to the model approach

***include appropriate references for each column

****specify if stress on development, leaf area expansion, stomatal control effecting carbon assimilation and/or transpiration, resource partitioning

2.3.2. *methods for evaluating model performance*

The models' relative performance in reproducing daily observed values of ET was assessed using multiple indicators. The coefficient of determination (R^2) was used to assess the degree of variation in the observed daily ET values that could be explained by the simulated values. A second ranking was based on the root mean square error (RMSE) between observed and simulated values determined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(ET_{sim,i} - ET_{obs,i}\right)^2} \quad (1)$$

where $ET_{sim,i}$ is the i^{th} simulation of daily ET, $ET_{obs,i}$ is the i^{th} observation of daily ET, and N is the total number of observation and simulation pairs. Both R^2 and RMSE were assessed using linear regression in R (R version 4.4.1) using RStudio (Posit team, 2024) with the R-package nlme for each environment and each environment by growth stage. After ranking the models, the values of the slopes and intercepts of regressions for each model and environment were plotted.

Additionally, RMSE was estimated across models and environments to explore relationships between model error and the evaporative demand. This was performed by grouping all observed and simulated pairs according to the value of the observed daily ET (0–2, 2–4, 4–6, 6–8, 8–10, 10–12 mm day^{-1}), growth stage and whether the models had correctly simulated the growth stage for that day. Similarly, relationships were explored between the RMSE in daily ET with RMSE in LAI simulations and simulations of total available soil water content.

The relationship between error in simulated ET and errors in simulated LAI and available soil water were explored in regression analysis for absolute mean error, mean errors, RMSE and the Root Mean Square Relative Error (RMSRE). The latter was estimated with Eq. (2) as:

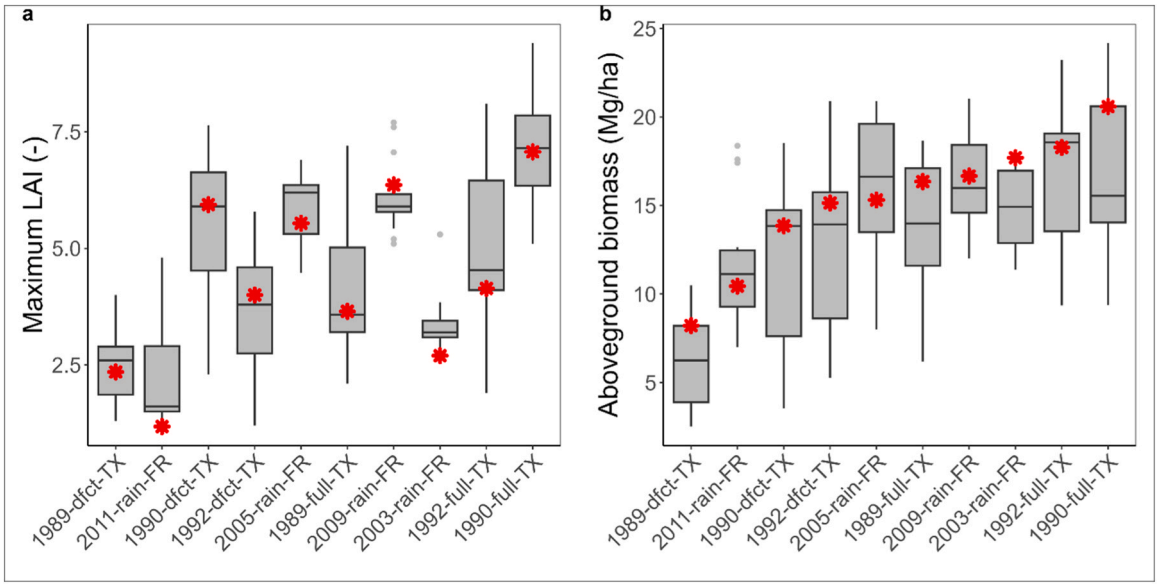

Fig. 2. Simulated values across all the model median (boxplots) by treatment. Observed values for each treatment displayed as red stars. Treatments are ranked from low to high observed maximum above ground biomass, as a proxy for crop stress. The boxplots show the 25th to 75th percentile of simulated values of the 17 models, with the medians indicated by the black lines in the boxes. Whiskers indicate 1.5 times the interquartile range (distance between the 25th and 75th percentile) for values lower than the 25th percentile and higher than the 75th percentile. Any additional outliers beyond these values are shown as individual points. The treatment name follows the convention: "sowing date year-water status-site".

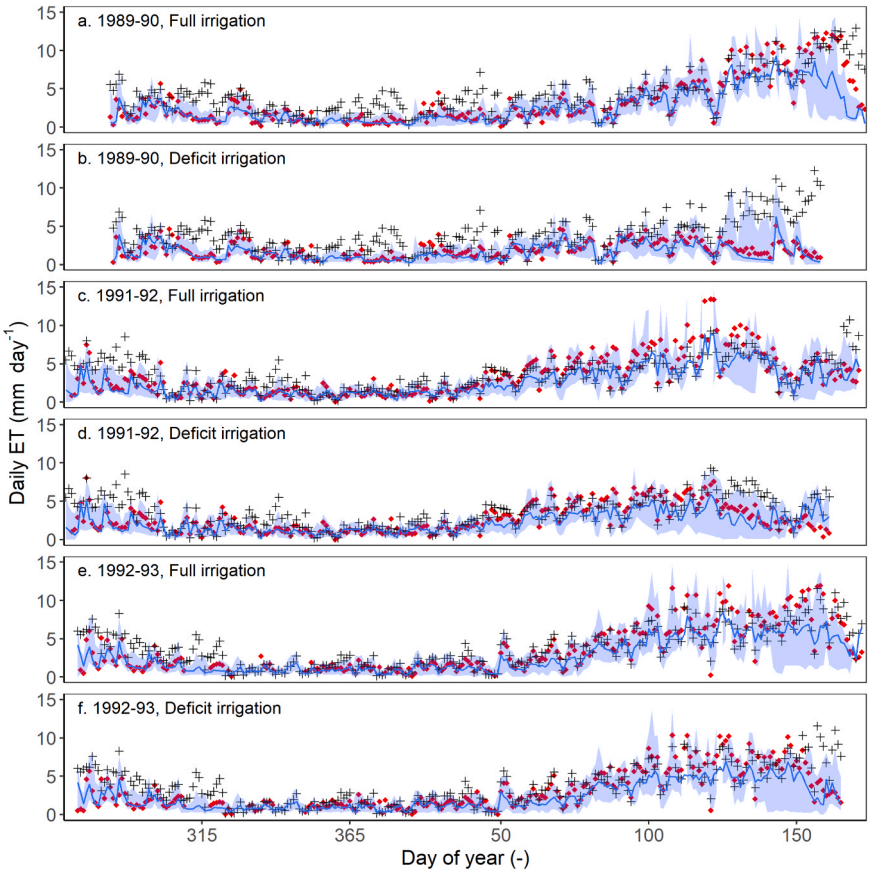

Fig. 3. Simulated and observed daily evapotranspiration (ET) over the (a, b) 1989–90, (c, d) 1991–92, and (e, f) 1992–93 growing seasons at bushland, texas, USA. Panels (a, c, e) are for fully irrigated treatments, while (b, d, f) indicate treatments with a partial irrigation application. Observed ET values from the weighing lysimeters are shown as red solid diamond shapes, estimates of daily reference crop evapotranspiration (ET_0) are indicated with black crosses, the median of simulated daily ET values from the model ensemble is shown as the blue line. The blue shaded region indicates the full range of values across the individual model simulations.

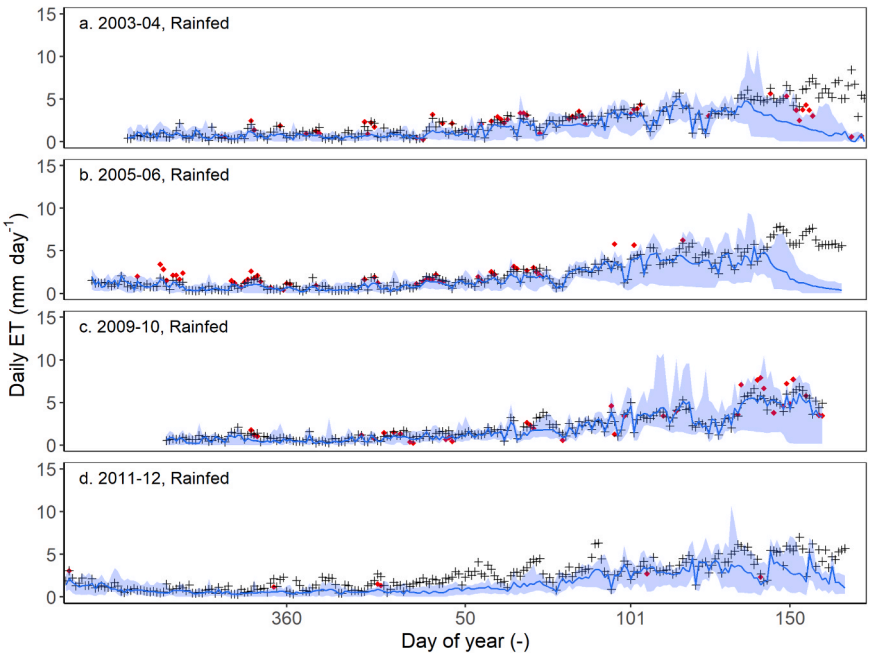

Fig. 4. Simulated and observed daily evapotranspiration (ET) over the (a) 2003–4, (b) 2005–6, (c) 2009–10, and (d) 2011–12 growing seasons growing seasons for avignon, France. Observed ET values from the eddy covariance tower are shown as red solid diamond shapes, simulated estimates of daily reference crop evapotranspiration (ET_0) are indicated with black crosses, the median of simulated daily ET values from the model ensemble is shown as the blue line. The blue shaded region indicates the full range of values across the individual model simulations.

$$RMSRE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{ET_{sim,i} - ET_{obs,i}}{ET_{obs,i}}\right)^2} \quad (2)$$

To test whether ET estimates of individual models explained more variation than ET_0 estimated from the FAO-56 Penman-Monteith (FAO-56 ET_0) combination equation for a grass reference crop (Allen et al., 1998), we first estimated the portion of variance in observed daily ET using linear models with FAO-56 ET_0 as the independent variable. Since the Bushland dataset had few missing values, we fitted a generalized least squares (GLS) regression model using the gls function from the nlme package in R. The model incorporated an autoregressive moving average (ARMA) correlation structure with two autoregressive terms (p = 2) to correct for serial correlation in the residuals, using date as the time variable. Missing values of observed ET were filled using the corresponding median estimate from all simulations. For Avignon, observed ET was modeled as a linear function of FAO-56 ET_0. For each location, the linear models were used to obtain separate estimates of the residual variance when FAO-56 ET_0 and simulated daily ET (ET_{sim}) were both included. For each combination of model and location, the portion of residual variance explained (PV) beyond that of the residual variance with only FAO-56 ET0 (RV_{FAO}) was then calculated as

$$PV = \frac{(RV_{FAO} - RV_{FAO+sim})}{RV_{FAO}} \quad (3)$$

where $RV_{FAO + sim}$ is the residual variance when both ET estimates were included in the linear model.

2.3.3. Sensitivity analysis calculations

The sensitivity analyses (Section 2.2.1) were assessed for three impact variables Y (E, T, and ET) for each model individually and for the model median for each experimental treatment using a one-way multivariate analysis of variance (ANOVA). For each impact variable (Y in Eq. 3), to determine the variance explained by each respective component (X_i, Table 2) as a portion of the total variability in the simulations, the main effects index (TS_i) was calculated for each X_i following (Monod et al., 2006) as:

$$TS_i = 1 - \frac{\mathrm{var}(E[Y|X_{-i}])}{\mathrm{var}(Y)} \quad (4)$$

3. Results

3.1. Observed and simulated summary growth variables and daily evapotranspiration

The end-of-season above-ground biomass reached over 15 Mg ha^{-1} in all irrigated treatments in Bushland and for most seasons in Avignon under rainfed conditions (Fig. 2). For Bushland, the 1989 sowing year season under deficit irrigation had very low biomass (8 Mg ha^{-1}), while at Avignon, the 2011 sowing produced the lowest biomass yields (10.5 Mg ha^{-1}). The models were largely able to reproduce the end-of-season biomass and maximum LAI values (Fig. 2). The model median was always close to the observed LAI, and for all experimental treatments, the model median captured well the seasonal LAI dynamics (SI Figs. S2 and S3). While the median also estimated end-season above-ground biomass reasonably well, there was a tendency to underestimate in many environments, particularly at higher biomass levels. This was unexpected as models mostly do not capture all possible growth limiting factors. Despite having access to the full data calibration set, some individual models exhibited very large errors (primarily underestimating) in simulation of biomass, as indicated by the large range of simulated values.

The time courses of daily ET for Bushland and Avignon are shown in Figs. 3 and 4, respectively. There were many more available observations for Bushland than Avignon because the quality controls for the eddy covariance system excluded many dates, whereas daily values were available for Bushland. The figures also show the full range of simulated daily ET across the full model ensemble, which is generally very large. Many of the observed ET values at both sites are greater than from any simulation, and they are considerably larger than the model ensemble, particularly for fully irrigated treatments. However, for the 1989–90 deficit irrigation treatment in Bushland (Fig. 3b), the model median seemed to largely capture the dynamics of water deficit later in the

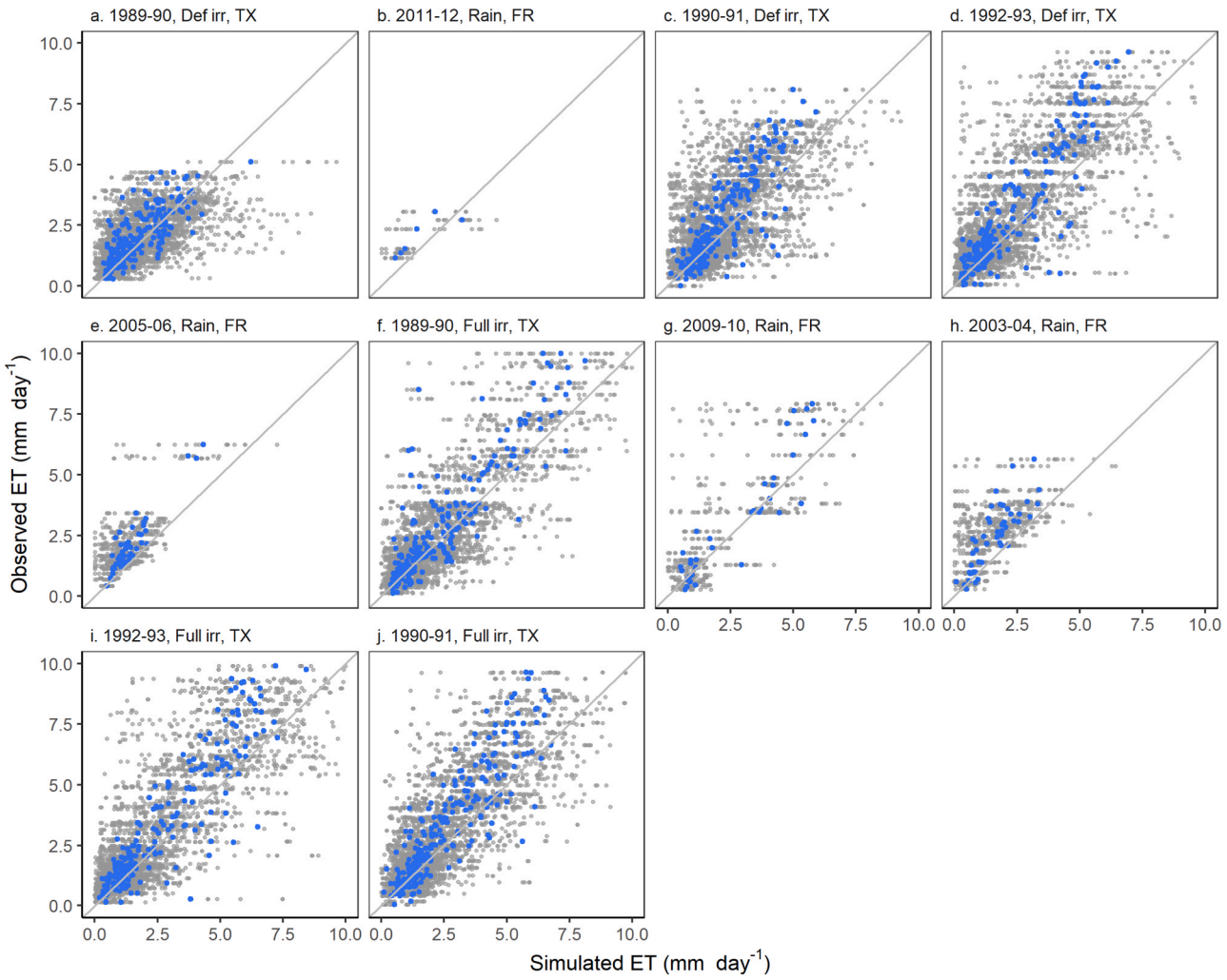

Fig. 5. Observed vs simulated daily ET values (mm day^{-1}), including data from all individual models and split by treatment. Dark blue data indicate the daily model ensemble median value, while individual daily model simulated values are shown in grey. Each panel is for a unique treatment (a) 1989–90, Deficit irrigation, TX, (b) 2011–12, Rainfed, FR, (c) 1990–91, Deficit irrigation, TX, (d) 1992–93, Deficit irrigation, TX, (e) 2005–06, Rainfed, FR, (f) 1989–90, Full irrigation, TX, (g) 2009–10, Rainfed, FR, (h) 2003–04, Rainfed, FR, (I) 1992–93, Full irrigation, TX, and (j) 1990–91, Full irrigation, TX. The panels are ordered from lowest to highest end-of-season biomass as a proximate indicator of how much water stress was experienced over the growing season.

season, though many models overestimated the actual reduced water use. Across all experimental treatments, there was a clear tendency for the model median, and therefore most models (SI Figs. S4 and S5), to underestimate water use (Fig. 5) with no clear pattern across different development stages (SI Fig S6).

3.2. *factors related to model skill in simulating crop water use*

As a first step to explore which factors explain model skill in simulating daily ET, the models were classified based on their use of an ET_0 versus directly applying an ET_P approach (Table 3). When ranked according to above-ground biomass across all environments, there was little evidence of any difference between the two approaches to simulate evaporative demand and water use (Fig. 6). This was largely consistent across environments, and the ranking was almost exactly conserved whether based on R^2 or RMSE. In fact, the top-ranking models were often the same DSSAT variant, implementing an ET_0 and then an ET_P approach, with other DSSAT variants having very low rankings with both their respective ET_0 and ET_P approaches. This was consistent across growth stages (data not shown). The slopes and intercepts of the top-ranking models tended to be closer to 1 and 0, respectively, than lower-ranking models (SI Fig S7).

We next examined whether the model error varied with atmospheric evaporative demand, as indicated by the observed daily ET values and the growth stage. Errors increased as the evaporative demand increased across all growth stages (Fig. 7). However, the relative error was approximately constant across levels of evaporative demand (data not shown). This error is primarily associated with an underestimation of ET (Fig. 5), though not for all cases, particularly models with lower R^2 values (Fig. 6 and SI Figs. S4 and S5). Further, evidence of this underestimation at levels of atmospheric evaporative demand is apparent for all growth stages (SI Fig. S6). That the error is largely consistent between the growth stages suggests that the error might be more related to evaporative demand than to partitioning per se (Fig. 7). Indeed, we saw that all models were able to explain more variation in the observed ET than the FAO-56 ET_0 estimate in all environments (SI Fig S8). This highlights that ET depends not only on atmospheric demand but also on soil water availability and leaf area. There was some evidence that error increased when models had not simulated the growth stage correctly for evaporative demands of 2–6 mm day^{-1}, whereas correct simulation of growth stage could not explain errors when evaporative demand was large (Fig. 7). This is supported by the relationship between mean error and RMSRE in simulated daily ET and simulated LAI during senescence (Fig. 8a, d) where an overestimation of LAI resulted in overestimation of

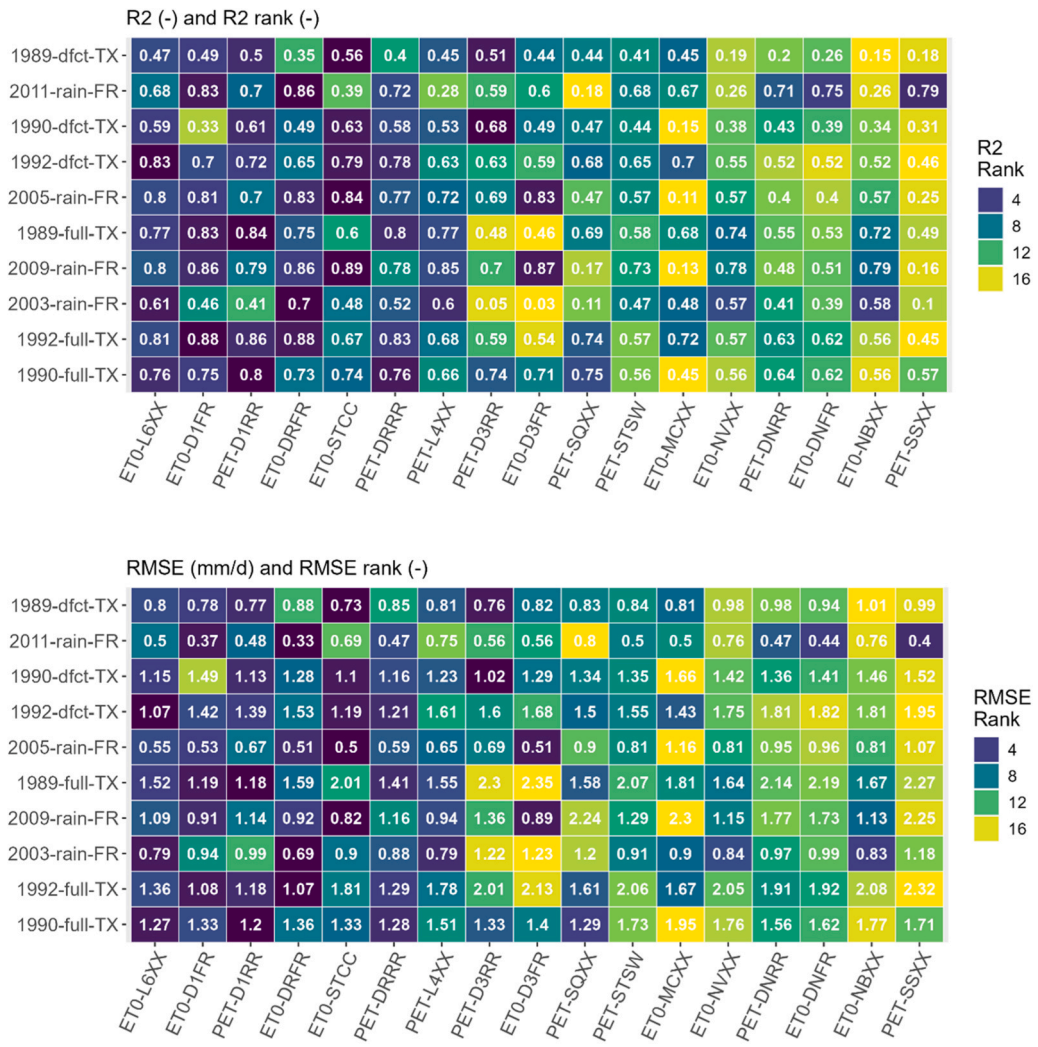

Fig. 6. Relative model performance ranking in terms of correlations (R^2, top panel, numbers) and root mean squared error (RMSE, bottom panel, numbers) for daily ET across the ten experimental treatments in bushland, texas (TX) and avignon, France (FR). each row represents one experimental treatment (ordered from highest observed above-ground biomass on bottom to lowest on the top). the treatment name follows the convention: sowing date year-water status-site. the water status indicates either deficit irrigation (dfct), full irrigation (full) or rainfed (rain). the 17 models are ordered by high to low correlation (R^2) across treatments (I.e., pooled values) for both panels. Model names are defined in table 1. the prefix to each model name indicates whether the model uses a reference evapotranspiration (ET_0) or a potential evapotranspiration (ET_P) approach. The colour of individual squares indicates each model's rank for that environment as compared to other models, with 1 indicating either the highest R^2 or lowest RMSE, respectively. Dark blue boxes indicate the top rank, and yellow boxes indicate lowest.

ET. The relationship between simulated ET and LAI errors at either early growth or senescence phase was very weak. There was weak support that errors in simulated soil water content in the root zone during the vegetative growth stage (when roots are nearer to the surface) were associated with larger errors in daily simulated ET (Fig. 8b, c). There was no relationship between errors in simulated soil water and errors in simulated daily ET during full cover or senescence growth stages (Fig. 9).

3.3. Factors and processes explaining uncertainty simulating crop water use

The results of the sensitivity analyses are presented for the median of the model ensembles for models using an ET_0 and ET_P approach separately in the top and bottom rows of Fig. 10, respectively. For models with an ET_0 approach, simulation of E was sensitive to methods used to simulate: (1) atmospheric demand, (2) partitioning between soil water evaporation and transpiration, and (3) soil water evaporation (Fig. 10a). Each component explained nearly equal shares of simulated variation, depending on the environment. The importance of partitioning of ET and how E was simulated varied across environments, though there was no clear pattern. For their simulation of crop T, demand and partitioning terms dominated (ET_0 and ET_P), with terms related to roots contributing little to variation for the studied environments (Fig. 10b). The method for simulation of soil water evaporation only explained a significant part of variation in the driest environment. Interaction terms, indicated by the residual, explained a large share of variation in some of the more water-limiting environments. Finally, for simulation of ET, process sensitivities largely reflected the pattern for T (Fig. 10c). For ET, the importance of the demand (ET_0) term generally increased for the least water-limiting environments. These patterns were generally maintained when looking at the sensitivity analysis results for the individual models implementing an ET_0 approach (SI Fig. S9). The greatest discrepancy among the models was in simulating T, where the relative importance of how atmospheric demand, partitioning of E and T and soil water evaporation varied between models, particularly for the more water-limiting environment (Fig. 11). For simulation of ET, there was large agreement between models on the processes describing the greatest share of variation, particularly for the least water-limiting conditions.

When considering the model median of the group of models not

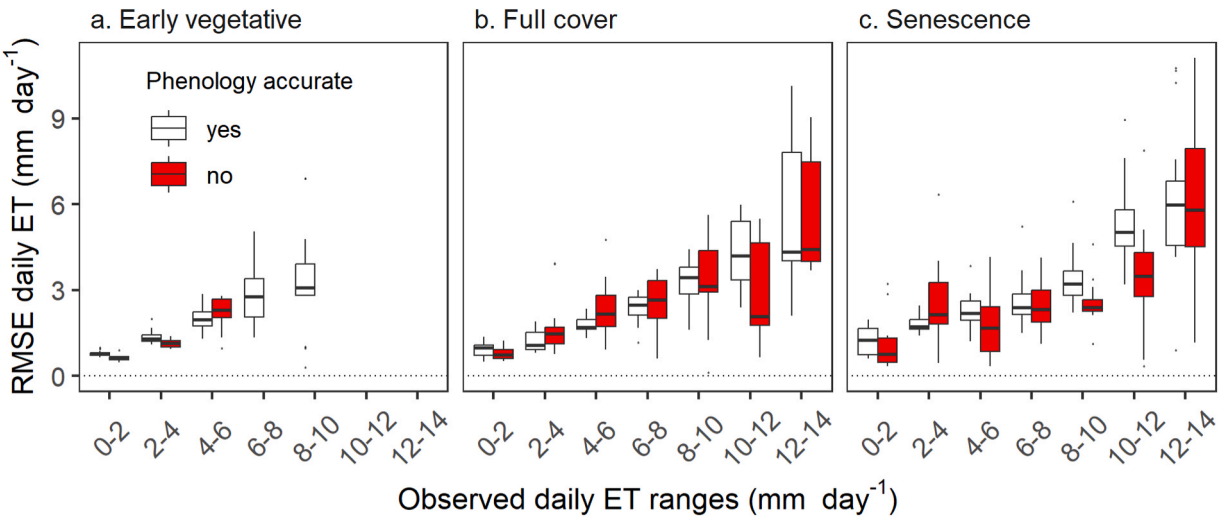

Fig. 7. Root mean squared error (RMSE) between observed and simulated daily ET values as a function of the observed daily ET values. The boxplots show the 25th to 75th percentile values across all experimental treatments and models. Top (bottom) whiskers extend to the minimum (maximum) of the maximum (minimum) value or the 75th (25th) percentile value plus (minus) 1.5 times the difference between the 75th and 25th percentile values. Data in the first panel is for the early vegetative phenological phase (soil water evaporation dominates); the middle panel is during full cover phenological stage (crop transpiration dominates) and the last panel shows data for the senescence phase. The phases were defined based on the LAI of the model median, which reasonably captured the observations. In each panel, White bars indicate RMSE for models which correctly simulated the growth stage at a particular date, while red bars correspond to the RMSE of model simulations which did not correctly simulate the phenological growth stage at a particular date.

using an ET_0 approach, both the atmospheric demand term (ET_P) and the soil water E terms explained near equal shares of uncertainty in simulation of E (Fig. 10d). This was true across all environments, except for the least water limiting case (1990, full irrigation, Bushland), in which the assumed maximum rooting depth also explained approximately one-third of total variation. For crop T, the demand term was not always the most important term explaining variation in simulations (Fig. 10e). While there was no clear pattern across environments, in some cases, the assumed rate of root depth increment mattered most, whereas in others the rate of soil water evaporation dominated. These results derive from very different behavior across the individual models (SI Figs. S10). For ET, the demand term (ET_P) explained the greatest share of variation in most environments, except for the most water-limited and the 2005 season for Avignon and the 1989 season in Texas with full irrigation (Fig. 10f). Unlike the models using the ET_0 approach, there was a large discrepancy in terms of the share of variation explained by the demand term versus soil water evaporation (Fig. 11).

The total amount of variation in the three explored impact variables was greater for the models using an ET_0 approach than those with an ET_P approach (SI Tables S2 and S3). For both model groups, total variation tended to increase with less water-limited environments. The most extreme case was an eight-fold increase in variation for the ET_0 model simulations of ET compared to treatments with full irrigation (SI Table S2).

4. Discussion

4.1. Underestimation of daily evapotranspiration

The foremost aim of the study was to evaluate the skill of wheat crop models' simulation of daily water use across a range of semi-arid and Mediterranean climatic conditions. This was motivated by the potentially important role crop models play in evaluating risk management and adaptation options, including irrigation water and infrastructure requirements (Elliott et al., 2014; Webber et al., 2016), to extreme weather events like drought stress. We detected a largely systematic bias in which the model median often underestimated water use. In agreement with the results of (Kang et al., 2009) using only three models, it did so consistently across environments and individual models largely did the same. While some models did overestimate water use, this was not as a systematic bias, but rather for models where simulations exhibited little correlation with observed ET.

While the relative error in underestimating daily ET was constant across levels of evaporative demand, the absolute error was greater for larger evaporative demand. This implies that errors in the soil water balance will increase more rapidly under warmer conditions, assuming no change in water supply, resulting in greater errors under future climate scenarios. While previous studies have demonstrated the large uncertainty in simulated water use by wheat crop models (Cammarano et al., 2016), this study is one of the first to quantify the error against observations and clearly demonstrates a consistent model underestimation. Interestingly, studies examining maize crop model simulations of water use did not detect a systematic bias across their model ensemble (Kimball et al., 2019, 2023).

The result is troubling as it suggests crop model-based climate change projections for wheat may be overly optimistic, underestimating irrigation water demand or drought stress and the associated yield losses in rainfed systems. Indeed, underestimation of water use when accumulated over many days results in simulated soil water availability being greater than measured in the field. In turn, crop models simulated a delay in depleted soil water content and failed to simulate reduced crop water use resulting from stomal closure as roots encounter drying soils. Indeed, we saw evidence of this in the simulations where crop water use was overestimated across the ensemble when observed water use was very low for the deficit irrigation treatments in Texas in 1989–90. Other studies at larger scales have detected possible overestimation of water use in extreme conditions by hydrological models, as detected in the 2003 heatwave across Europe (Schewe et al., 2019) or in climate change projections (Milly and Dunne, 2017). Milly and Dunne (2016) explain the overestimation associated as being due to their reliance on potential ET where reduced ET with water limitation is not accounted for (Milly and Dunne, 2017).

4.2. Factors related to model skill in simulating crop water use

We also sought to identify which factors or processes lead to the greatest error or uncertainty in simulating daily ET to prioritize model improvement needs. We initially intended to accomplish this with a

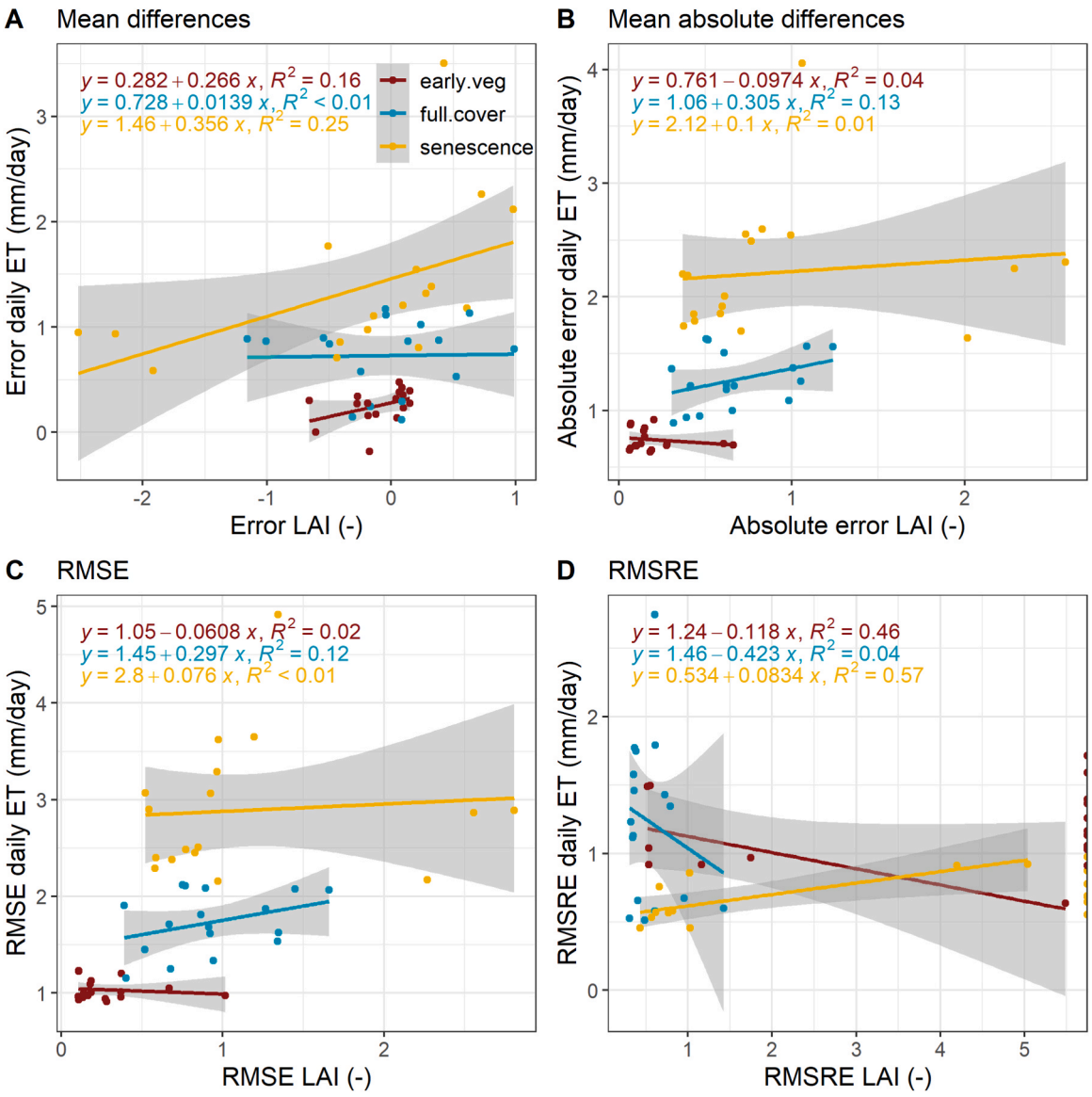

Fig. 8. Comparisons of the error between observed and simulated daily ET values on the y-axis and error in LAI on the x-axis with errors estimated as (A) mean differences, (B) absolute differences, (C) root mean square error (RMSE), and (D) root mean square relative error (RMSRE). errors were calculated for each model and phenological development stage, pooled across all environments, though data by model are not distinguished in the figure. Colours indicate early vegetative (red), full cover (aqua blue) and senescence (yellow) development stages. The phases were defined based on the LAI of the model median, which reasonably captured the observations.

sensitivity analysis. Our starting hypothesis was that for environments characterized by full irrigation, the demand-related terms would explain the large share of variation in simulations, while root growth-related terms would predominantly explain variation in more water-limited environments. The hypothesis that there would be a process by environment interaction was not supported by the analysis. The demand terms (ET_0 and ET_P), and to a lesser extent soil water evaporation simulation method, explained most of the variation in simulations of the model median. This result is in agreement with the studies by Kimball et al. (2019) and Kimball et al. (2023) for maize models and studies at global levels looking at hydrological cycles (Kingston et al., 2009). There were exceptions to this for some of the models implementing an ET_P approach, where the assumed maximum rooting depth or rate of root depth increment explained large shares of variation in simulated transpiration and ET (SI Fig. S10). Similarly, Webber et al. (2016) found that root traits explained large shares of variation in seasonal ET in dry environments, but very little in fully irrigated conditions. That the root-related terms did not explain a large share of variation in simulated water use is considered here as good news, as it is difficult to obtain quality data for model development and calibration of below-ground growth and root traits (Kersebaum et al., 2015). However, this requires more study including soil water content dynamics which is beyond the scope of the current study. Indeed, most of the crop models considered in this study have simple approaches to simulate root growth, soil organic matter and crop water uptake dynamics that do not match the state of the art (Tougma et al., 2025), though model complexity depends on the research question.

As the sensitivity analysis identified the importance of demand terms, we then explored if models using an ET_0 approach performed better than those directly using an ET_P approach. We did not explicitly test this, but the ranking of model performance in terms of R^2 or RMSE supplied no evidence that one group of models performed better than another. Here we can especially consider the four DSSAT models, which each conducted simulations with both FAO56-ET_0 (Allen et al., 1998) and Priestley Taylor ET_P (Priestley and Taylor, 1972) approaches. For these models, the approach to simulating ET did not influence their

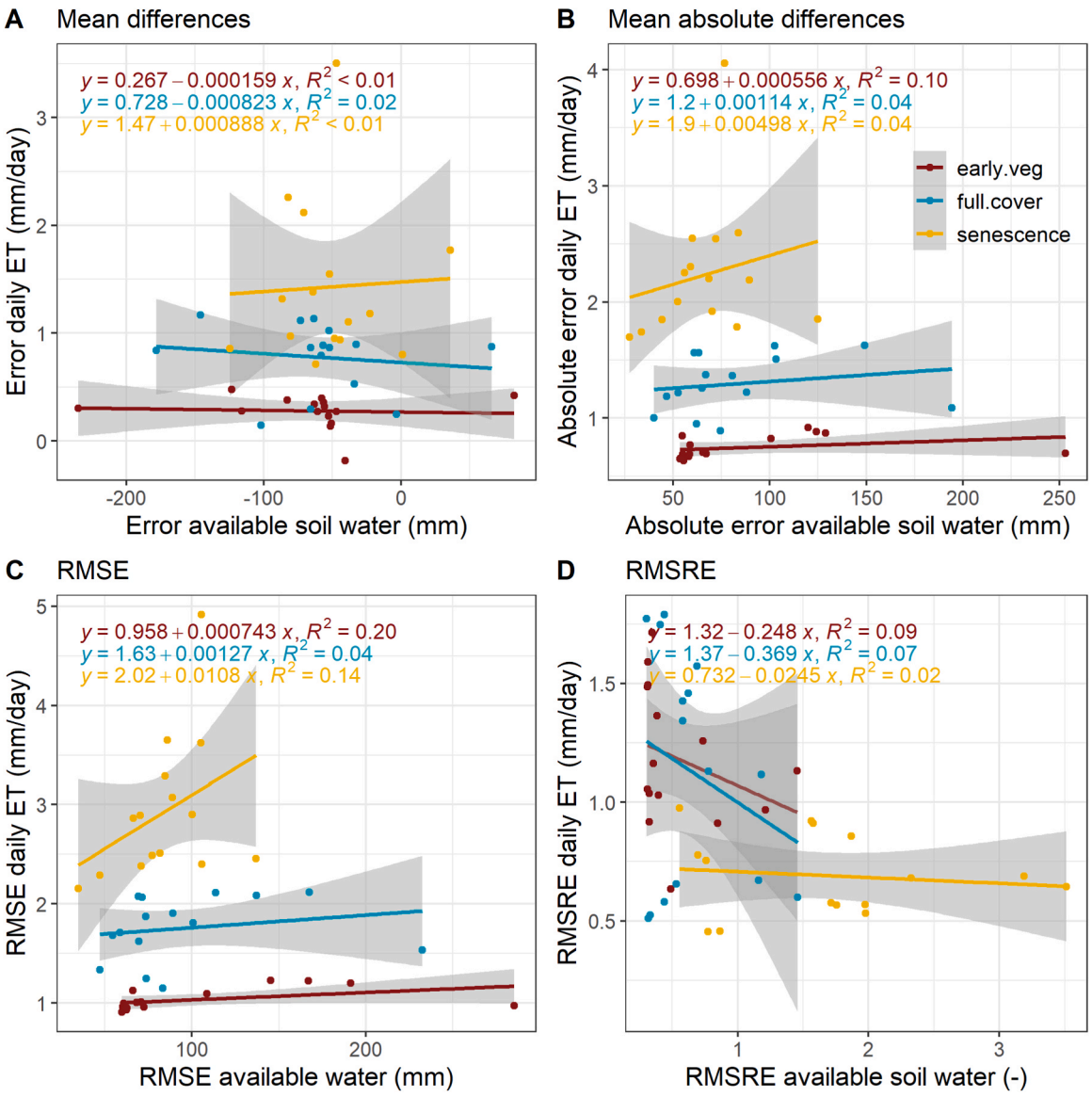

Fig. 9. Comparisons of the error between observed and simulated daily ET values on the y-axis and errors in cumulative seasonal available soil water in the soil profile on the x-axis with errors estimated as (A) mean differences, (B) absolute differences, (C) root mean square error (RMSE), and (D) root mean square relative error (RMSRE). errors were calculated for each model and phenological development stage, pooled across all environments, though data by model are not distinguished in the figure. For the site in avignon, available soil water was considered to a depth of 150 cm, while for the bushland, texas site a depth of 200 cm was considered. Rmse was calculated for each model and phenological development stage, pooled across all environments, though data by model are not distinguished in the figure. Colours indicate early vegetative (red), full cover (aqua blue) and senescence (yellow) development stages. The phases were defined based on the LAI of the model median, which reasonably captured the observations.

ranking. Rather, the two versions of the same DSSAT model variant tended to be very closely ranked, irrespective of the approach to simulate ET.

There was limited evidence that errors in the simulation of ET were related to the error in simulating LAI. Studies at global level identified differences in simulated LAI or ground cover as the main factor describing the underestimation of T/ET in global earth system models (Lian et al., 2018; Villegas et al., 2015). However, this effect became only visible when the crop was in the full cover stage, which is dominated by crop transpiration. However, we saw no evidence that this error was related primarily to partitioning, as errors between the three growth stages were approximately equal (Fig. 7), and the regression between daily ET and LAI errors was not significant in the early vegetative phases (Fig. 8).

4.3. *Limitations of study and future research*

This study is the first to test several wheat crop models against observational data for their ability to simulate daily ET. While the study uses observations from two excellent experiments, it covers a limited range of production conditions and semi-arid agroecological conditions. It is unclear how generalizable the results are for global crop production. The exercise should be repeated for both temperate and tropical environments, where drought risk is rather periodic, with conditions that are dominated by high humidity during frequent rainy periods. While drought in these conditions may have similar atmospheric conditions as those explored in the current study, water use at low vapour pressure deficit must be correctly simulated for satisfactory simulations of the soil water balance. Further, elevated atmospheric CO_2 concentrations (e $[CO_2]$) with climate change will also affect transpiration through both increasing transpiration efficiency and increasing LAI (Kimball, 2016)

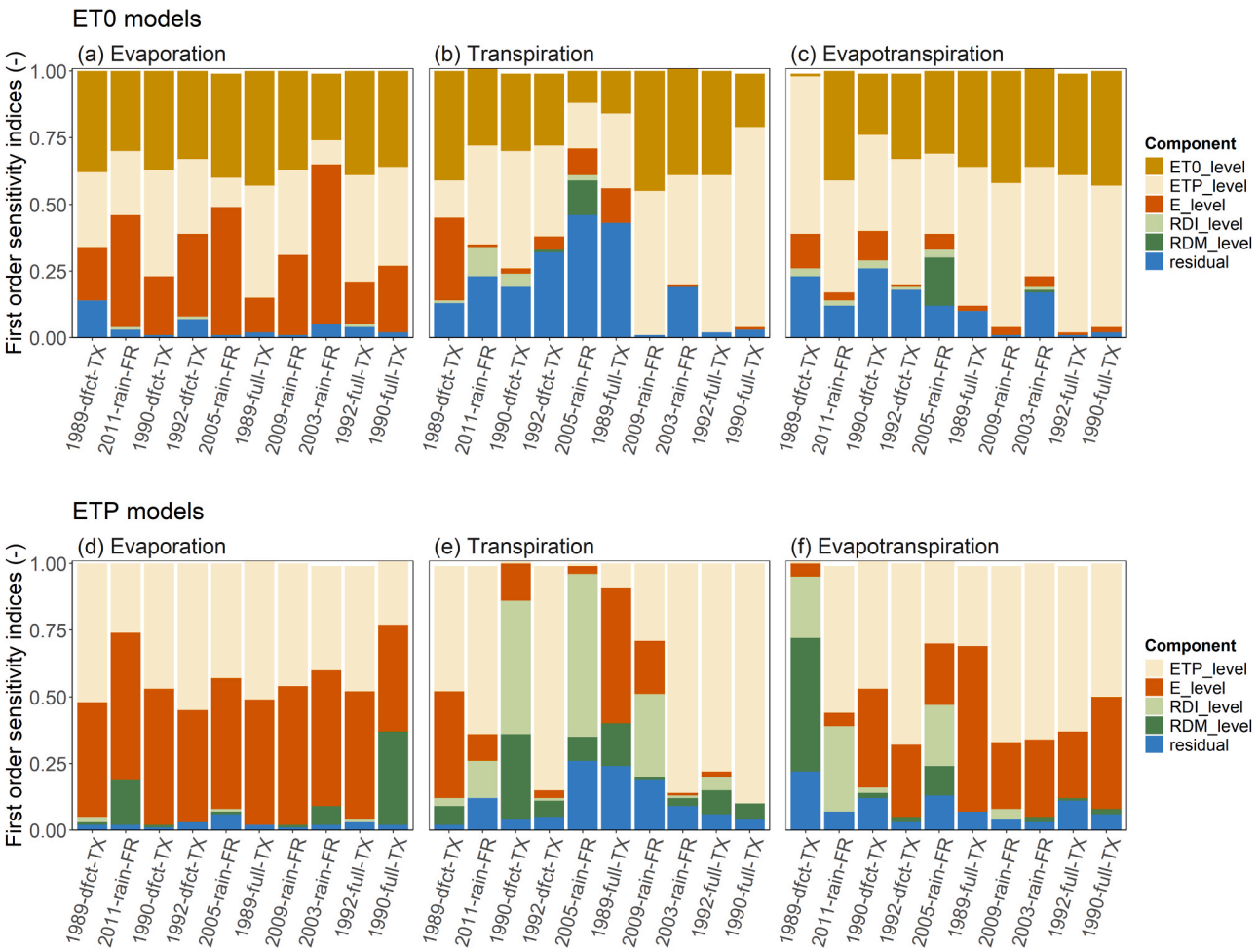

Fig. 10. First-order sensitivity index of the model median for three seasonal cumulative impact variables: soil water evaporation (a, d), crop transpiration (b, e) and crop evapotranspiration (c, f) for each environment. Relative variation explained by each component (source of variation) is indicated with coloured bars for ET_0 (brown), ET_P (cream), e (Orange), daily increment in root depth, RDI (light Green), maximum root depth, RDM (dark Green) and residuals (blue). panels (a), (b), and (c) show the median of the models implementing reference crop evapotranspiration (ET_0) approach, while panels (d), (e), and (f) show the median of models using a potential evapotranspiration (ET_P) approach. Environments are named following the convention: sowing date year-water status-site and are ordered from lowest to high observed maximum above-ground biomass as a proxy of the degree of water stress.

with considerable uncertainty about their combined effects on wheat water use (Bourgault et al., 2018; Wang et al., 2020). Our current study did not address water use under e[CO_2], and this should be prioritized for future study.

Unfortunately, the Avignon dataset lacked many observations of daily ET, which introduces the possibility that our assessment is biased towards environmental conditions that lead to closure of the Eddy Covariance energy balance. While the filtered dataset ensured that ET values were of high quality, the incomplete time series precluded estimating total seasonal ET. Similarly, while the datasets from both locations contained frequent measurements of LAI, they were insufficient to delimit daily LAI dynamics, that could be used to define growth stages (see *Section 2.3.1*). As such, we relied on model median LAI values, introducing further uncertainty. Results of errors with increasing atmospheric evaporative demand (Fig. 7) are sensitive to how one defines the growth stages, and this was simply a qualified assumption due to the limited number of LAI observations.

A further concern is that the apparent consistent underestimation of ET might partially arise from bias in the measurements of ET or incorrect characterization of the soil profiles. For example, if the wheat crops were able to extract soil moisture below the specified permanent wilting point, this might increase actual T relative to what was simulated. Similarly, if actual drainage was slower than the modeled soil profile allowed, actual ET might exceed simulated ET following irrigations or large precipitation events. Such issues might be partially resolved by examining measured vs. simulated soil moisture in the deeper soil layers, which will be undertaken in a subsequent study. Quality control of the Bushland lysimeter. weather, and soil water content data was described in the Methods and Materials section. Questions about how well the Bushland weighing lysimeter data represent ET in the surrounding fields were explored for cotton crops by Evett et al. (2012) and further explored for maize and sorghum crops by Evett et al. (2019). No consistent under- or over-estimation of ET by the lysimeters was found except for one lysimeter with cotton where LAI on the lysimeter was appreciably different from LAI in the surrounding field. Differences in wheat cover between the lysimeters and the fields in the three wheat seasons at Bushland were not observed, leading us to believe that the lysimeter ET reported was representative of the surrounding fields.

While this study focused on daily ET, model skill in simulating soil water dynamics throughout the profile is critical to identify drought adaptative traits and to understand soil water retention, groundwater recharge as well as leaching of nitrate and agrochemicals to the groundwater. Further, given the limited availability of data on soil water use or crop ET, it will be important to quantify model performance with more limited model calibration than undertaken here, particularly given that evaluation of root growth and soil water extraction simulations is

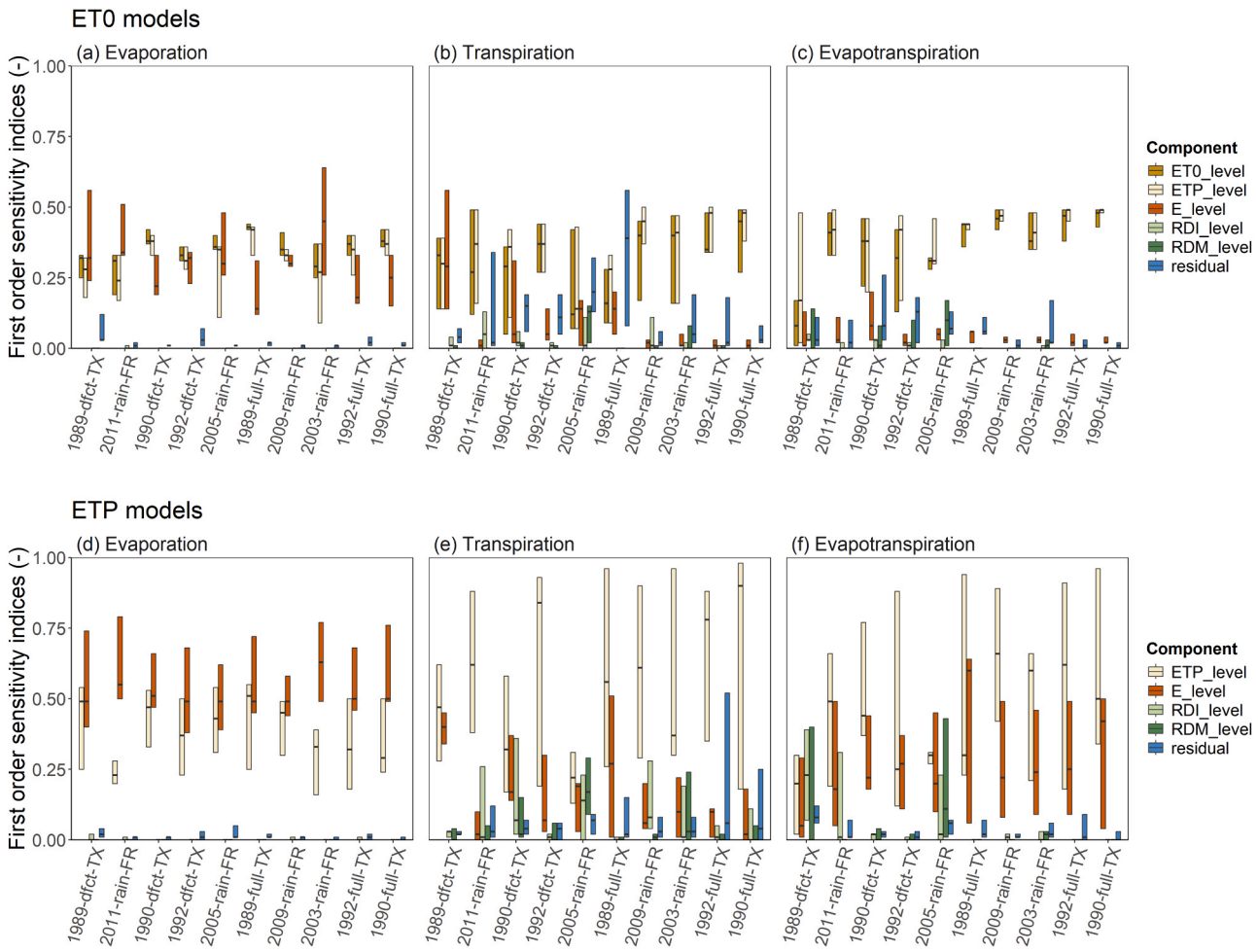

Fig. 11. The range across models of the first-order sensitivity index for three seasonal cumulative impact variables: soil water evaporation (a, d), crop transpiration (b, e) and crop evapotranspiration (c, f) for each environment. Relative variation explained by each component (source of variation) is indicated with coloured bars for ET_0 (brown), ET_P (cream), e (Orange), daily increment in root depth, RDI (light Green), maximum root depth, RDM (dark Green) and residuals (blue). panels (a), (b), and (c) show the model median for models implementing a reference crop evapotranspiration (ET_0) approach, while panels (d), (e), and (f) show the median of models using a potential evapotranspiration (ET_P) approach. Environments are named following the convention: sowing date year-water status-site and are ordered from lowest to high observed maximum above-ground biomass as a proxy of the degree of water stress.

very limited as compared to above ground growth.

5. Conclusions

In investigating the performance of an ensemble of wheat crop models to simulate daily water dynamics for semi-arid and Mediterranean environments, we found the model median underestimated daily water use. This underestimation increased at high evaporative demand. Much of this error was related to underestimating atmospheric demand and errors in simulating LAI, at least when the crop was at full cover and actively transpiring. Despite the importance of correctly simulating demand, our results did not detect that either one of two major approaches to simulate demand (ET_P or ET_0) was superior to the other. The results call for a thorough investigation of the reasons why crop models generally underestimate ET and then further testing the accuracy of simulated ET, particularly environments where production conditions are riskier varying between humid periods of rainfall to hot dry periods. It may be that in other rainfed environments where water is more limiting that errors in simulation of leaf area or root growth will increase the errors in water use and water balance. Our results imply that crop-model-based climate change impact studies underestimate the negative impacts of future warming driving drought stress and yield losses in wheat.

CRediT authorship contribution statement

H. Webber: Conceptualized study, designed study, prepared data, analysed results, prepared figures, wrote the paper, coordinated the study and provided the funding. **D. Cooke**: Prepared data, analysed results, prepared figures, wrote the paper. **C. Wang:** Analysed results, prepared tables, reviewed the paper. **S. Asseng**: Discussed concept, conducted simulations, reviewed the paper. **P. Martre**: Discussed concept, conducted simulations, reviewed the paper. **F. Ewert**: Discussed concept, reviewed the paper. **B. Kimball**: Discussed concept, reviewed the paper. **G. Hoogenboom**: Conducted simulations, reviewed the paper. **S. Evett**: Discussed concept, provided data, reviewed the paper. **A. Chanzy**: Discussed concept, provided data, reviewed the paper. **S. Garrigues**: Provided data, reviewed the paper. **A. Olioso**: Provided data, reviewed the paper. **K.S. Copeland**: Provided data, reviewed the paper. **J.L. Steiner**: Provided data, reviewed the paper. **D. Cammarano**: Conducted simulations, reviewed the paper. **M. Crépeau**: Conducted simulations, reviewed the paper. **E. Diamantopoulos**: Conducted simulations, reviewed the paper. **R. Ferrise**: Conducted simulations, reviewed the paper. **L. Manceau**: Conducted simulations, reviewed the paper. **T. Gaiser**: Conducted simulations, reviewed the paper. **Y. Gao**: Conducted simulations, reviewed the paper. **S. Gayler**: Conducted simulations, reviewed the paper. **J.R. Guarin**: Conducted

simulations, reviewed the paper. **T. Hunt**: Conducted simulations, reviewed the paper. **G. Jégo**: Conducted simulations, reviewed the paper. **G. Padovan**: Conducted simulations, reviewed the paper. **E. Pattey**: Conducted simulations, reviewed the paper. **D. Ripoche**: Conducted simulations, reviewed the paper. **A. Rodríguez**: Conducted simulations, reviewed the paper. **M. Ruiz-Ramos**: Conducted simulations, reviewed the paper. **V. Shelia**: Conducted simulations, reviewed the paper. **A.K. Srivastava**: Conducted simulations, reviewed the paper. **I. Supit**: Conducted simulations, reviewed the paper. **F. Tao**: Conducted simulations, reviewed the paper. **K. Thorp**: Conducted simulations, reviewed the paper. **M. Viswanathan**: Conducted simulations, reviewed the paper. **T. Weber**: Conducted simulations, reviewed the paper. **J. White**: Conceptualized study, designed study, prepared data, analysed results, prepared figures, wrote the paper.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We gratefully acknowledge the contributions of the late Terry A. Howell, Sr., whose leadership in establishing and leading the USDA ARS large weighing lysimeter program at Bushland, Texas, USA was instrumental in developing crop growth and water use data for not only winter wheat, but many other crops grown in the semi-arid U.S. Southern High Plains. HW and CW were funded by the Leibniz Female Professorship Award, Germany (Application number: P102/2020). HW and DC acknowledge funding by the German Research Foundation, Germany (DFG Project Number: 470400637). SE and KS were funded by the USDA ARS Ogallala Aquifer Program, U.S.A.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.fcr.2025.110032.

Data availability

The protocols and all data used in this study are published and available at: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/HDKKAL.

References

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. FAO irrigation and drainage paper no. 56. Rome Food Agric. Organ. U. Nations 56 (97), e156.
Asseng, S., et al., 2004. Simulated wheat growth affected by rising temperature, increased water deficit and elevated atmospheric CO2. Field Crops Res. 85 (2), 85–102.
Asseng, S., et al., 2013. Uncertainty in simulating wheat yields under climate change. Nat. Clim. Change 3 (9), 827–832.
Bourgault, M. et al., 2018. Will Early Vigour Still be Useful in Wheat in a Future with Elevated Atmospheric CO 2?, ASA, CSSA, and CSA International Annual Meeting (2018). ASA-CSSA-SSSA.
Brisson, N., Itier, B., L'Hotel, J.C., Lorendeau, J.Y., 1998. Parameterisation of the Shuttleworth-Wallace model to estimate daily maximum transpiration for use in crop models. Ecol. Model. 107 (2-3), 159–169.
Brisson, N., Perrier, A., 1991. A semiempirical model of bare soil evaporation for crop simulation models. Water Resour. Res. 27 (5), 719–727.
Brisson, N., Seguin, B., Bertuzzi, P., 1992. Agrometeorological soil water balance for crop simulation models. Agric. For. Meteorol. 59 (3-4), 267–287.
Bruckler, L., Lafolie, F., Doussan, C., Bussières, F., 2004. Modeling soil-root water transport with non-uniform water supply and heterogeneous root distribution. Plant Soil 260 (1), 205–224.
Cammarano, D., et al., 2016. Uncertainty of wheat water use: simulated patterns and sensitivity to temperature and CO2. Field Crops Res. 198, 80–92.
Elliott, J., et al., 2014. Constraints and potentials of future irrigation water availability on agricultural production under climate change. Proc. Natl. Acad. Sci. 111 (9), 3239–3244.
Enders, A., et al., 2023. SIMPLACE—a versatile modelling and simulation framework for sustainable crops and agroecosystems. silico Plants 5 (1), diad006.
Evett, S.R., et al., 2016. The bushland weighing lysimeters: a quarter century of crop ET investigations to advance sustainable irrigation. Trans. ASABE 59 (1), 163–179.
Evett, S.R., et al., 2020. Past, present, and future of irrigation on the US great plains. Trans. ASABE 63 (3), 703–729.
Evett, S. et al., 2022a. The Bushland, Texas, Winter Wheat Datasets.
Evett, S., Heng, L., Moutonnet, P., Nguyen, M., 2008. Field estimation of soil water content: a practical guide to methods, instrumentation, and sensor technology. IAEA, Vienna.
Evett, S., Howell, T., Schneider, A., Copeland, K., Dusek, D., 1994. Energy and water balance modeling of winter wheat. Am. Soc. Agric. Engr. Pap. (942022): 49085-9639.
Evett, S.R., Marek, G.W., Colaizzi, P.D., Brauer, D.K., O'Shaughnessy, S.A., 2019. Corn and sorghum ET, E, yield, and CWP as affected by irrigation application method: SDI versus mid-elevation spray irrigation. Trans. ASABE 62 (5), 1377–1393.
Evett, S.R., Marek, G.W., Colaizzi, P.D., Copeland, K.S., Ruthardt, B.B., 2022b. Methods for downhole soil water sensor calibration—Complications of bulk density and water content variations. Vadose Zone J. 21 (6), e20235.
Evett, S.R., Marek, G.W., Colaizzi, P.D., Copeland, K.S. and Ruthardt, B.B., 2024. Spreadsheet for lysimeter data analysis, Bushland, Texas.
Evett, S.R., Marek, G.W., Copeland, K.S., Colaizzi, P.D., 2018. Quality management for research weather data: USDA-ARS, bushland, TX. Agrosyst. Geosci. Environ. 1 (1), 1–18.
Evett, S.R., Schwartz, R.C., Howell, T.A., Baumhardt, R.L., Copeland, K.S., 2012. Can weighing lysimeter ET represent surrounding field ET well enough to test flux station measurements of daily and sub-daily ET? Adv. Water Resour. 50, 79–90.
Evett, S., Steiner, J., 1995. Precision of neutron scattering and capacitance type soil water content gauges from field calibration. Soil Sci. Soc. Am. J. 59 (4), 961–968.
Evett, S., Tolk, J., Howell, T., 2003. A depth control stand for improved accuracy with the neutron probe. Vadose Zone J. 2 (4), 642–649.
Feddes, R., 1978. 4.3 Simulation of field water use and crop yield.
Gabaldón-Leal, C., et al., 2016. Modelling the impact of heat stress on maize yield formation. Field Crops Res. 198, 226–237.
Garrigues, S., et al., 2015. Evaluation of land surface model simulations of evapotranspiration over a 12-year crop succession: impact of soil hydraulic and vegetation properties. Hydrol. Earth Syst. Sci. 19 (7), 3109–3131.
Gayler, S., et al., 2013. Assessing the relevance of subsurface processes for the simulation of evapotranspiration and soil moisture dynamics with CLM3. 5: comparison with field data and crop model simulations. Environ. Earth Sci. 69, 415–427.
Goudriaan, J. (1977). Crop micrometeorology: a simulation study. Wageningen University and Research.
Haxeltine, A., Prentice, I.C., 1996. BIOME3: an equilibrium terrestrial biosphere model based on ecophysiological constraints, resource availability, and competition among plant functional types. Glob. Biogeochem. Cycles 10 (4), 693–709.
Howell, T., et al., 1993. Radiation balance and soil water evaporation of bare Pullman clay loam soil, management of irrigation and drainage systems: integrated perspectives. ASCE, pp. 922–929.
Howell, T., et al., 2006. Crop coefficients developed at bushland, texas for corn, wheat, sorghum, soybean, cotton, and alfalfa. World Environ. Water Resour. Congr. 2006 Examining Conflu. Environ. Water Concerns 1–9.
Howell, T., Schneider, A., Dusek, D., Marek, T., Steiner, J., 1995a. Calibration and scale performance of bushland weighing lysimeters. Trans. ASAE 38 (4), 1019–1024.
Howell, T., Steiner, J., Schneider, A., Evett, S., 1995b. Evapotranspiration of irrigated winter wheat—Southern high plains. Trans. ASAE 38 (3), 745–759.
Huang, J., et al., 2017. Dryland climate change: recent progress and challenges. Rev. Geophys. 55 (3), 719–778.
IPCC, 2022. Climate change 2022: impacts, adaptation, and vulnerability. Contribution of working group II to the sixth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, UK and New York, NY, USA, p. 3056.
Jamieson, P., Francis, G., Wilson, D., Martin, R., 1995. Effects of water deficits on evapotranspiration from barley. Agric. For. Meteorol. 76 (1), 41–58.
Jones, J.W., et al., 2003. The DSSAT cropping system model. Eur. J. Agron. 18 (3-4), 235–265.
Jones, A.C., Kiniry, J.R. and Dyke, P., 1986. CERES-Maize: A simulation model of maize growth and development.
Kang, S., Payne, W.A., Evett, S.R., Robinson, C.A., Stewart, B.A., 2009. Simulation of winter wheat evapotranspiration in texas and henan using three models of differing complexity. Agric. Water Manag. 96 (1), 167–178.
Kersebaum, K.C., et al., 2015. Analysis and classification of data sets for calibration and validation of agro-ecosystem models. Environ. Model. Softw. 72, 402–417.
Kimball, B.A., 2016. Crop responses to elevated CO2 and interactions with H2O, N, and temperature. Curr. Opin. Plant Biol. 31, 36–43.
Kimball, B.A., et al., 2019. Simulation of maize evapotranspiration: an inter-comparison among 29 maize models. Agric. For. Meteorol. 271, 264–284.
Kimball, B.A., et al., 2023. Simulation of evapotranspiration and yield of maize: an inter-comparison among 41 maize models. Agric. For. Meteorol. 333, 109396.
Kingston, D.G., Todd, M.C., Taylor, R.G., Thompson, J.R., Arnell, N.W., 2009. Uncertainty in the estimation of potential evapotranspiration under climate change. Geophys. Res. Lett. 36 (20).
Koppa, A., et al., 2024. Dryland self-expansion enabled by land–atmosphere feedbacks. Science 385 (6712), 967–972.
Lesk, C., Rowhani, P., Ramankutty, N., 2016. Influence of extreme weather disasters on global crop production. Nature 529 (7584), 84–87.

Lian, X., et al., 2018. Partitioning global land evapotranspiration using CMIP5 models constrained by observations. Nat. Clim. Change 8 (7), 640–646.
Liu, Y., Kumar, M., Katul, G.G., Feng, X., Konings, A.G., 2020. Plant hydraulics accentuates the effect of atmospheric moisture stress on transpiration. Nat. Clim. Change 10 (7), 691–695.
Maiorano, A., et al., 2017. Crop model improvement reduces the uncertainty of the response to temperature of multi-model ensembles. Field Crop Res 202, 5–20.
Marek, G.W., et al., 2014. Post-processing techniques for reducing errors in weighing lysimeter evapotranspiration (ET) datasets. Trans. ASABE 57 (2), 499–515.
Mauder, M., et al., 2013. A strategy for quality and uncertainty assessment of long-term eddy-covariance measurements. Agric. For. Meteorol. 169, 122–135.
Milly, P.C.D., Dunne, K.A., 2016. Potential evapotranspiration and continental drying. Nat. Clim. Change 6 (10), 946–949.
Milly, P.C.D., Dunne, K.A., 2017. A hydrologic drying bias in Water-Resource impact analyses of anthropogenic climate change. JAWRA J. Am. Water Resour. Assoc. 53 (4), 822–838.
Monod, H., Naud, C., Makowski, D., 2006. Uncertainty and sensitivity analysis for crop models. Work. Dyn. Crop Model. Eval. Anal. Parameter Appl. 4, 55–100.
Peng, B., et al., 2020. Towards a multiscale crop modelling framework for climate change adaptation assessment. Nat. Plants 6 (4), 338–348.
Penman, H.L., 1948. Natural evaporation from open water, bare soil and grass. Proc. R. Soc. Lond. Ser. A. Math. Phys. Sci. 193 (1032), 120–145.
Priesack, E., Gayler, S., Hartmann, H.P., 2006. The impact of crop growth sub-model choice on simulated water and nitrogen balances. Nutr. Cycl. Agroecosyst. 75, 1–13.
Priestley, C.H.B., Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Mon. Weather Rev. 100 (2), 81–92.
Rezaei, E.E., et al., 2023. Climate change impacts on crop yields. Nat. Rev. Earth Environ. 4 (12), 831–846.
Ritchie, J.T., 1972. Model for predicting evaporation from a row crop with incomplete cover. Water Resour. Res. 8 (5), 1204–1213.
Ritchie, J.T., 1998. Soil water balance and plant water stress. Underst. Options Agric. Prod. 41–54.
Ritchie, J., Godwin, D., Otter-Nacke, S., 1985. CERES-wheat: a user-oriented wheat yield model. Preliminary documentation. AGRISTARS Publication No, YM-U3-04442-JSC-18892. Michigan State University, Michigan.
Ritchie, J.T., Porter, C.H., Judge, J., Jones, J.W., Suleiman, A.A., 2009. Extension of an existing model for soil water evaporation and redistribution under high water content conditions. Soil Sci. Soc. Am. J. 73 (3), 792–801.
Rosenzweig, C., et al., 2013. The agricultural model intercomparison and improvement project (AgMIP): protocols and pilot studies. Agric. For. Meteorol. 170, 166–182.
Schewe, J., et al., 2019. State-of-the-art global models underestimate impacts from climate extremes. Nat. Commun. 10 (1), 1005.
Semenov, M.A., Martre, P., Jamieson, P.D., 2009. Quantifying effects of simple wheat traits on yield in water-limited environments using a modelling approach. Agric. For. Meteorol. 149 (6-7), 1095–1104.
Stocker, B.D., et al., 2019. Drought impacts on terrestrial primary production underestimated by satellite monitoring. Nat. Geosci. 12 (4), 264–270.
Tao, F., Zhang, Z., Liu, J., Yokozawa, M., 2009. Modelling the impacts of weather and climate variability on crop productivity over a large area: a new super-ensemble-based probabilistic projection. Agric. For. Meteorol. 149 (8), 1266–1278.
Tao, F., Zhang, Z., 2013. Climate change, wheat productivity and water use in the north China plain: a new super-ensemble-based probabilistic projection. Agric. For. Meteorol. 170, 146–165.
Tougma, I.A., et al., 2025. AMPSOM: a measureable pool soil organic carbon and nitrogen model for arable cropping systems. Environ. Model. Softw. 185, 106291.
Van Ittersum, M.K., Rabbinge, R., 1997. Concepts in production ecology for analysis and quantification of agricultural input-output combinations. Field Crops Res. 52 (3), 197–208.
Villegas, J.C., et al., 2015. Sensitivity of regional evapotranspiration partitioning to variation in woody plant cover: insights from experimental dryland tree mosaics. Glob. Ecol. Biogeogr. 24 (9), 1040–1048.
Wang, E.L., et al., 2017. The uncertainty of crop yield projections is reduced by improved temperature response functions. Nat. Plants 3 (10), 833-833.
Wang, S., et al., 2020. Recent global decline of CO2 fertilization effects on vegetation photosynthesis. Science 370 (6522), 1295–1300.
Webber, H., et al., 2016. Uncertainty in future irrigation water demand and risk of crop failure for maize in Europe. Environ. Res. Lett. 11 (7), 074007.
Webber, H., et al., 2017. Canopy temperature for simulation of heat stress in irrigated wheat in a semi-arid environment: a multi-model comparison. Field Crops Res. 202, 21–35.
Webber, H., et al., 2018. Physical robustness of canopy temperature models for crop heat stress simulation across environments and production conditions. Field Crops Res. 216, 75–88.
Webber, H., et al., 2020. No perfect storm for crop yield failure in Germany. Environ. Res. Lett. 15 (10), 104012.
Webber, H., Rezaei, E.E., Ryo, M., Ewert, F., 2022. Framework to guide modeling single and multiple abiotic stresses in arable crops. Agric. Ecosyst. Environ. 340, 108179.
Weber, T.K., Durner, W., Streck, T., Diamantopoulos, E., 2019. A modular framework for modeling unsaturated soil hydraulic properties over the full moisture range. Water Resour. Res. 55 (6), 4994–5011.
Weber, T.K., Finkel, M., da Conceição Gonçalves, M., Vereecken, H., Diamantopoulos, E., 2020. Pedotransfer function for the Brunswick soil hydraulic property model and comparison to the van Genuchten-Mualem model. Water Resour. Res. 56 (9), e2019WR026820.
Wheeler, T., Von Braun, J., 2013. Climate change impacts on global food security. Science 341 (6145), 508–513.
White, J.W., et al., 2013. Integrated description of agricultural field experiments and production: the ICASA version 2.0 data standards. Comput. Electron. Agric. 96, 1–12.
White, J.W., Hoogenboom, G., Kimball, B.A., Wall, G.W., 2011. Methodologies for simulating impacts of climate change on crop production. Field Crops Res. 124 (3), 357–368.
Zhao, M., Liu, Y., Konings, A.G., 2022. Evapotranspiration frequently increases during droughts. Nat. Clim. Change 12 (11), 1024–1030.