Performance of a wheat yield prediction model and factors influencing the performance: A review and meta-analysis

Shirui Hao[1], Dongryeol Ryu[1], Andrew Western[1], Eileen Perry[2,1], Heye Bogena[3], Harrie Jan Hendricks Franssen[3]

[1] *Department of Infrastructure Engineering, Faculty of Engineering and Information Technology, The University of Melbourne*

[2] *Agriculture Victoria, Department of Jobs, Precincts and Regions*

[3] *Institute of Bio- and Geosciences, Agrosphere Institute (IBG-3), Forschungszentrum Jülich*

Abstract

CONTEXT: Process-based crop models provide ways to predict crop growth, evaluate environmental impacts on crops, test various crop management options, and guide crop breeding. They can be used to explore options for mitigating climate change impacts when combined with climate projections and explore mitigation of environmental impacts of production. The Agricultural Production Systems SIMulator (APSIM) is a widely adopted crop model that offers modules for simulation of various crops, soil processes, climate, and grazing within a modelling system that enables robust addition of new components.

OBJECTIVE: This study uses APSIM Classic-Wheat as an example to examine yield prediction accuracy of biophysically based crop yield modelling and to analyse the factors influencing the model performance.

METHODS: We analysed yield prediction results of APSIM Classic-Wheat from 76 published studies across thirteen countries on four continents. In addition, a meta-database of modelled and observed yields from 30 studies was established and used to identify factors that influence yield prediction uncertainty.

RESULTS AND CONCLUSIONS: Our analysis indicates that, with site-specific calibration, APSIM predicts yield with a root mean squared error (RMSE) smaller than 1 t/ha and a normalised RMSE (NRMSE) of about 28%, across a wide range of environmental conditions for independent evaluation periods. The results show increasing errors in yield with limited modelling information and adverse environmental conditions. Using soil hydraulic parameters derived from site-specific measurements and/or tuning cultivar parameters improves yield prediction accuracy: RMSE decreases from 1.25 t/ha to 0.64 t/ha and NRMSE from 32% to 14%. Lower model accuracy was found where APSIM overestimates yield under high water deficit condition and when it underestimates yield under nitrogen

limitation. APSIM severely over-predicts yield when some abiotic stresses such as heatwaves and frost affect the crop growth.

SIGNIFICANCE: This paper uses APSIM-Wheat as an example to provide perspectives on crop model yield prediction performance under different conditions covering a wide spectrum of management practices, and environments. The findings deepen the understanding of model uncertainty associated with different calibration processes or under various stressed conditions. The results also indicate the need to improve the model's predictive skill by filling functional gaps in the wheat simulations and by assimilating external observations (e.g., biomass information estimated by remote sensing) to adjust the model simulation for stressed crops.

Keywords: Cropping system, APSIM Classic, wheat, yield prediction performance, meta-analysis, literature review

1. Introduction

Biophysical models, as agricultural simulation systems, are widely used to simulate crop growth, test management options, assess environmental trade-offs, and explore ways to cope with climate change impacts. The key strength of process-based biophysical models is their embodiment of our understanding of the dynamic interactions among crop, soil, water, atmosphere and solar radiation within the agricultural system (Horie et al., 1992). In essence, they simulate the biological and physical processes linking environmental effects to crop yield outcomes (Roberts et al., 2017). These models can assist in quantifying the impacts of changing climate on crop yield, designing efficient management practices, and informing crop breeding to secure food production. But deficiencies in the models and their implementations (e.g., calibration and weather inputs) can introduce random or systematic errors leading to uncertain yield predictions. While current efforts are underway to improve biophysical schemes, model inputs and implementation, understanding the current state of process-based model performance and sources of uncertainty can guide us to more effective strategies.

There exist several widely used process-based crop models that include Agricultural Production System SIMulator (APSIM) (Brown et al., 2018; Holzworth et al., 2014; Keating et al., 2003; McCown et al., 1996, 1995), Simulateur mulTIdisciplinaire pour les Cultures Standard (STICS) (Brisson et al., 2003, 2002, 1998), Environmental Policy Integrated Climate (EPIC) (Williams et al., 1989), The Soil & Water Assessment Tool (SWAT) (Neitsch et al., 2011), Decision Support System for Agrotechnology Transfer (DSSAT) (Jones et al., 2003), WOrld FOod STudies (WOFOST) (Van Diepen et al., 1989; van Ittersum et al., 2003), Soil Water Atmosphere Plant (SWAP) (Van Dam et al., 1997) and AquaCrop (Hsiao et al., 2009; Raes et al., 2009; Steduto et al., 2009). This work focuses on APSIM Classic as an example to explore a biophysical model's performance in predicting yield and the factors influencing the performance.

APSIM has been used for research and practical applications globally for over 25 years. It is also available as an online commercial agricultural decision-support tool, named Yield Prophet®, to serve Australian growers (Carberry et al., 2009; Hochman et al., 2009b). APSIM consists of interconnected modules describing the biophysical roles of soil water, soil nutrients, organic matter, crops, weather, and management. It can simulate various crop types and pastures. Simulated crops include wheat (Asseng et al., 2000, 1998a), maize (Archontoulis et al., 2014; Shamudzarira and Robertson, 2002), canola (Robertson and Lilley, 2016) and various legumes (Robertson et al., 2002). Previous studies have used it as a tool to reproduce the biophysical processes of the cropping system from paddock to regional level (Araya et al., 2020; Gaydon et al., 2006; Keating et al., 2002), including representing the role of soils (Connolly et al., 2002; Probert and Dimes, 2004; Thorburn et al., 2001), the influence of climate (Asseng et al., 2015; Bahri et al., 2019), and animal grazing (Bosi et al., 2020; Holzworth et al., 2014). It has also been used to guide genotype design of future cultivars (Rötter et al., 2015) and to understand genotype, environment and management interactions (Casadebaig et al., 2016; Hammer et al., 2010; Manschadi et al., 2006; Martre et al., 2015a; Zheng et al., 2015). Researchers have also combined APSIM with various climate projection models to investigate future food security challenges and explore solutions to mitigate environmental impacts on production (Akinseye et al., 2020; Anwar et al., 2020; Asseng et al., 2011, 2004; Liu et al., 2016a; Ludwig and Asseng, 2006). It has been coupled with economic models to develop profit maximisation strategies and to study the effectiveness of crop insurance (Hansen et al., 2009; Van Wijk et al., 2014). As a cropping system tool, the accuracy and uncertainty of APSIM simulations under different environmental and input resources conditions are important to model users, as they need to be aware of the uncertainty in model outputs under the circumstances of their interest.

Globally, wheat is the fourth most-produced crop and provides 20% of the calories consumed by people (FAO, 2020; Shiferaw et al., 2013). APSIM-Wheat yield prediction accuracy has been extensively evaluated for research applications and as a decision support tool for farmers. In addition to evaluations of APSIM-Wheat at field or regional scales with particular management practices or wheat cultivars, several APSIM developers and researchers have also collected assessment datasets covering a broader spectrum of management practices, environments, and cultivars to analyse model strengths, weaknesses and identify aspects for further development. An extensive set of the model validation data and descriptions are available on the APSIM website (https://www.apsim.info/). Holzworth et al. (2011) presented part of the wheat final yield validation results from those datasets, reporting a coefficient of determination (R^2) of 0.93 and root mean squared error (RMSE) of 0.46 t/ha. Brown et al. (2014) compared the predicted against observed yields for 164 simulations under a wide range of environments and treatments, resulting in an R^2 =0.92. Gaydon et al. (2017) reviewed APSIM performance across various cropping systems in Asia and identified its strengths and weaknesses with 43 experimental datasets from 12 countries. They concluded that the model could be further improved in aspects related to harsh environments, conservation agriculture and low input systems. Brown et al. (2018) validated the model with experimental datasets from 8 countries covering a broad range of crop treatments. The results demonstrated that the model performed well overall with an $R^2 \geq 0.84$ and Nash-Sutcliffe Efficiency (NSE)≥0.81.

While extensive work has been done to evaluate the model yield prediction accuracy, factors that affect the model's yield prediction uncertainty remain to be investigated comprehensively. In general, model prediction uncertainty originates from deficient/inaccurate model structure, input forcing data, parameter specification and observations used for model calibration/validation (Vrugt et al., 2008). In this paper, we review and quantify APSIM Classic (which hereafter is referred to as "APSIM")-Wheat yield prediction accuracy by compiling existing evaluation datasets from the literature and analysing the contribution of environmental and input resource factors to the model prediction uncertainty. The objective of the study is to review the performance of process-based crop model yield prediction and identify influential factors affecting prediction accuracy, with APSIM-Wheat used as an example. Firstly, an overview of the APSIM-Wheat yield prediction accuracy and uncertainty is provided by collating the model evaluation results from published studies. Next, a meta-analysis based on existing literature is performed to identify the factors influencing uncertain yield prediction, which include model specification and calibration, heat and frost stresses, water and nitrogen availability. The uncertainties in yield prediction associated with the above-mentioned factors are discussed. Finally, suggestions are provided for improving the accuracy of crop models such as APSIM-Wheat prediction under circumstances of high prediction uncertainty.

2. Review of APSIM-Wheat model evaluation

2.1. Overview of the APSIM Classic and Wheat module

APSIM is an agricultural modelling platform equipped with various biophysical and management modules to simulate cropping systems (Holzworth et al., 2014; Keating et al., 2003). The model is composed of multiple modules that simulate soil water, nutrients (carbon, nitrogen, and phosphorus), and crop growth processes under different environmental and management conditions. For example, the SoilWat (Jones and Kiniry, 1986; Littleboy et al., 1992) calculates soil water movement using a cascading water balance model, and it is used by most APSIM users (all studies reviewed in this work used SoilWat). Soil Water Infiltration and Movement (SWIM) is another option to simulate the soil-water-solute balance based on Richards' equation and the advection-dispersion equation, but is not adopted by most model users. The SoilN module simulates the transformations of carbon and nitrogen in the soil. SoilWat and SoilN interact with each other and together provide plant available soil water and nitrogen information to the Wheat module (Zheng et al., 2014) for simulating crop growth. The Wheat module simulates phenological development, plant morphology, biomass and nitrogen concentration of different wheat components, grain number and grain size on a daily basis (Keating et al., 2001). Here we use APSIM-Wheat to collectively represent the wheat growth simulation model which consists of the required APSIM modules including SoilWat, SoilN, and Wheat. A detailed description of the Wheat module is provided by Zheng et al. (2014). We only provide an overview of the stress factors considered in Wheat since they are used to better understand the factors influencing yield prediction performance.

Water stress: The Wheat module accounts for water stress impacts in simulating photosynthesis and leaf expansion. The influence on photosynthesis (f_{W_photo}) and leaf expansion (f_{W_expan}) is calculated as follows:

$$f_{W\_photo} = \frac{W_u}{W_d}, \quad (1)$$

$$f_{W\_expan} = h_{w\_expan} \times \frac{W_u}{W_d}, \quad (2)$$

where W_u and W_d are crop water uptake and water demand, respectively, and h_{w_expan} is a water stress factor piecewise linearly related to W_u/W_d. Smaller W_u/W_d results in a smaller h_{w_expan} value. So, equation (2) is effectively a quadratic function of W_u/W_d. Equations (1) and (2) indicate that both biomass accumulation and leaf expansion are scaled by the ratio of total daily water uptake to crop water demand, with leaf expansion more sensitive to the water stress.

Nitrogen stress: The Wheat module accounts for nitrogen stress on phenology (not applied), biomass accumulation, leaf appearance and expansion, and grain filling. The stress for these aspects is determined by the difference between organ nitrogen concentration and minimum and critical nitrogen concentration.

Heat stress: The Wheat module takes temperature as a factor affecting the crop into account in many ways (Zheng et al., 2014). The daily maximum temperature is considered as the temperature stress in calculating LAI senescence. The daily mean temperature $(T_{max} + T_{min})/2$ is considered as the stress factor affecting wheat growth in (1) crop phenology via the thermal time; (2) root depth growth; (3) biomass accumulation; (4) biomass demand of grain and the rate of grain filling.

Frost stress: The Wheat module incorporates the leaf area senescence effect using a frost stress function; however, the default parameterisation of the stress factor results in zero impact during the whole simulation, which means it is not in application.

2.2. Literature search and selection criteria

We performed a literature search for peer-reviewed journal articles focused on APSIM-Wheat performance evaluation using Scopus, ISI Web of Science and Google Scholar. The following keywords in English were employed to search the literature: APSIM, wheat, Triticum aestivum, yield prediction, validation, evaluation, verification, and performance. A total of 108 articles published between September 1997 and February 2020 are reviewed. Among these, only the 76 articles that included independent validation datasets (independent growing seasons/fields from calibration) of APSIM-Wheat grain yield prediction using *in situ* yield data at field scale are used for the meta-analysis of APSIM-Wheat yield uncertainty. The APSIM-Wheat validation datasets from these papers are across thirteen countries in four continents, including Australia (41 studies), New Zealand (2 studies), United States of America (1 study), Belgium (1 study), The Netherland (1 study), Turkey (1 study), China (20 studies), India (3 studies), Pakistan (2 studies), Syria (1 study), Iran (2 studies), Ethiopia (3 studies), Tunisia (1 study) (some papers include locations from several countries, Figure 1). The evaluation sites cover a broad range of environmental and management conditions such as extreme temperatures, different water and nutrient availability situations, various soil types and hydraulic conditions.

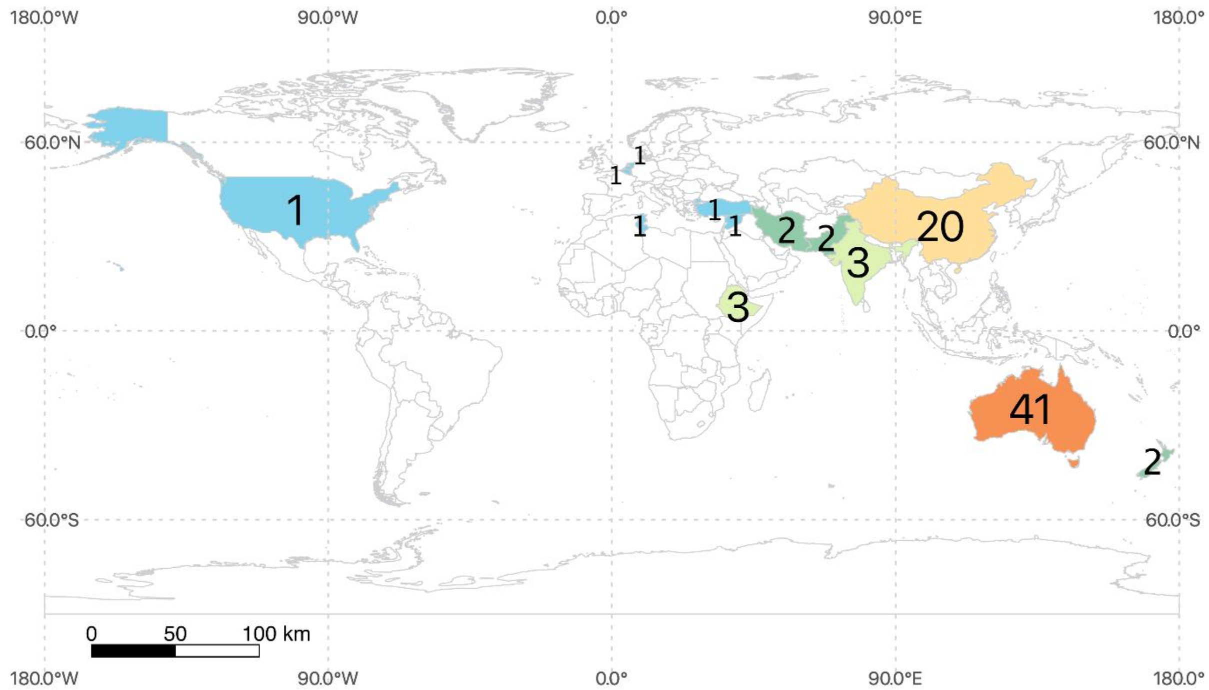

Figure 1. Number of articles for each country (the dataset of United States of America is in the conterminous United States)

2.3. APSIM-Wheat calibration and evaluation metrics

The model evaluation datasets in reviewed papers contain calibration and validation processes. Here calibration refers to all processes to improve the model fit to data, while validation refers to testing models against independent data not used in calibration to ensure the rigour of the model evaluation. In model calibration, variables that are related to crop growth, such as physiological dates, leaf area index (LAI), biomass, yield or soil water content and evapotranspiration are typically considered as the benchmarks for calibration and validation. Based on different data sources used, three calibration (or parameter setting) methods were defined in this paper: (1) Manual/automatic tuning of parameters to make the model simulations better fit the observations; (2) Direct specification of parameters using field measurements of these parameters; (3) Parameter specification using available databases (e.g., APSoil) or estimated data such as estimating lower limits from soil texture. The first two methods are collectively referred to as a fully site-specific calibration. If only one of them is adopted, it is partially site-specific calibration. The third method is classified as non-site-specific calibration (*Table 1*).

Table 1. Calibration methods defined in this paper

Manual tuning of parameters	Site-specific calibration
Parameter specification using ground observations.	
Parameter specification using APSoil or estimated data.	Non-site-specific calibration

Many researchers specify the specific cultivar used in the simulation or manually adjust genetic parameters, especially those controlling wheat phenology and yield development by trial-and-error to improve the model predictions against field observations. The genetic parameters used to characterise the cultivar are summarised in *Table 2*. The reported calibrated values of these parameters are summarised in Supplementary Table S1. Some coefficients listed in Supplementary Table S1 were derived from results for multiple soil types, sowing dates, sites, and growing seasons, which should help ensure the model robustness.

Table 2. Definition of the genetic parameters

Generic parameter name	Unit	Definition
tt_end_of_juverile	°C	The thermal time from end of juvenile to terminal spikelet stage
tt_floral_initiation	°C	The thermal time target for floral initiation
tt_flowering	°C	The thermal time target for flowering
tt_start_grain_fill	°C	The thermal time target to start grain filling stage
tt_end_grain_fill	°C	The thermal time target to end grain filling stage
tt_startgf_to_mat	°C	The thermal time target from beginning of grain filling to maturity
potential_grain_filling_rate	g/(grain °Cd)	Potential grain filling rate
grain_per_gram_stem	grain	Numbers of grain per gram stem
max_grain_size	g	Maximum grain size
vern_sens	N/A	Sensitivities to vernalisation
photop_sens	N/A	Sensitivities to photoperiod
phyllochron	°Cd	Phyllochron interval

Soil parameters such as soil texture, soil hydraulic, and chemical parameters were usually specified in studies using laboratory test data (soil samples were taken from study fields), APSoil soil database (Dalgliesh et al., 2012, 2009), semblable objects or estimated data, such as estimating lower limits from soil texture (Sadras et al., 2003).

Several statistical criteria are commonly selected to evaluate model performance: the coefficient of determination (R^2), root mean square error (*RMSE*, also referred to as root mean square difference, *RMSD*), normalised *RMSE* (*NRMSE*), model efficiency (*EF*), and/or index of agreement (*d*) defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2}, \tag{3}$$

$$NRMSE = RMSE/\overline{O}, \tag{4}$$

$$EF = 1 - \frac{\sum_{i=1}^{N}(P_i - O_i)^2}{\sum_{i=1}^{N}\left(O_i - \overline{O}\right)^2}, \tag{5}$$

$$d = 1 - \frac{\sum_{i=1}^{N}(P_i - O_i)^2}{\sum_{i=1}^{N}\left(\left|P_i - \overline{O}\right| + \left|O_i - \overline{O}\right|\right)^2}, \tag{6}$$

where P_i and O_i represent *i*th predicted and observed values, respectively, $\overline{O}$ the mean observed values, and N the sample size. R^2 measures the goodness-of-fit of a linear relationship between simulated and observed values, and hence ignores model bias. R^2 is also sensitive to the variance of the samples. *RMSE* and *NRMSE* represent the mean difference of predictions and observations, and they include measures of both bias and random errors. *EF* and *d* assess the degree of model prediction and are similar to R^2, except they are influenced by both bias and random errors. The index of agreement *d* is normalised by a measure of combined spread in observations and predictions, while *EF* (and R^2) are normalised by the spread in observations. The model reproduces experimental data perfectly when R^2=1, *RMSE*=0, *NRMSE*=0, *EF*=1 and *d*=1.

2.4. APSIM-Wheat yield prediction performance

Table *3* presents basic information on each paper validation datasets – reference, APSIM version, location, year, cultivar, environmental conditions, treatments, APSIM performance, and main model application. All reviewed works used APSIM Classic (version 1.X – version 7.9). The model has been applied mostly at plot or paddock, and sometimes regional, scales as a cropping system tool solely to assess the environmental impacts on food production, or combined with other models (e.g., climate projection models, economic models) to investigate future food security challenges and explore solutions or to develop profit maximisation strategies and study the effectiveness of crop insurance. A full version of Table 3 with detailed information is shown in Supplementary Table S2.

Table 3. List of validation datasets from the literature used in this study (P(gs): growing season rainfall; P(yr): average annual rainfall; P(t): total annual rainfall; T(gs): growing season temperature; Tg(yr): average annual temperature; Tx: annual maximum temperature; Tn: annal minimum temperature; SWC: soil water content; ESW: extractable soil water; ET: evapotranspiration; N: nitrogen fertiliser; DC31: wheat growth stage code, stem elongation; DC65: anthesis; * Data were used to compose the meta-database for further analysis in Section 3)

Dataset No.	Reference	APSIM version	Location	Years	Cultivar	Environmental conditions	Treatments	APSIM performance	Application
1	(Probert et al., 1995)	1.X	Warwick, Queensland, Australia	1969-1992	Timgalen (1969-1974, 1978-1981), Cook (1983-1984), Kite (1985-1987), Hartog (1990), QT4118 (1992)	In some seasons, crops suffered from diseases like root-lesion nematodes and crown rot.	2 tillage managements: conventional, no tillage. 2 crop residue managements: stubble burned, retained. 3 N application rates: 0, 23, 69 kg N/ha	Yield-RMSE=0.937 t/ha, R^2=0.30.	Model development and validation
2	(Probert et al., 1998)	1.X	Gatton, Queensland, Australia	1992-1995	N/A	N/A	A range of nitrogen inputs and under different moisture regimes.	Predicted yield = 1.03 * observed yield - 0.27 (t/ha), R^2=0.78.	Model evaluation
3	(Asseng et al., 1998b)	NWheat	Beverley, Merredin, Moora, and Wongan Hills, Western Australia, Australia	Beverley (1990-1993), Merredin (1973, 1986), Moora (1994-1995), Wongan Hills (1983, 1994)	Dagger, Gamenya, Gutha, Kulin, Spear	Beverley: P(yr)=421 mm, P(gs)=352 mm, soil type: duplex. Merredin: P(yr)=310 mm, P(gs)=234 mm, soil type: duplex. Moora: P(yr)=458 mm, P(gs)=388 mm, soil type: deep sand. Wongan Hills: P(yr)=386 mm, P(gs)=318 mm, soil type: loamy sand.	Different nitrogen supply, irrigation, sowing date, sowing density, and deep ripping.	Observed yield range=1.0 to 4.0 t/ha, R^2=0.77, RMSD=0.4 t/ha. Observed biomass range=0.1 to 11.0 t/ha, R^2=0.90, RMSD=0.8 t/ha. Observed LAI range=0 to 3.8 m^2/m^2, R^2=0.59, RMSD=0.6 m^2/m^2.	Model evaluation
4	(Asseng et al., 1998a)	NWheat	Moora, Western Australia, Australia	1994-1995	Spear	Deep sand, P(yr)=459 mm, ranged from 203 to 790 mm.	N treatments of 0, 50, and 90 kg N/ha.	Discrepancies between observed and predicted yields are less than 0.4 t/ha.	Establish the probability of yield
5	(Asseng et al., 2000)	NWheat	The Eest, PAGV, The	1983-1984	Arminda	The Eest and PAGV: Soil type: silty loam,	N applications: The Eest: 0, 60, 110, 150, 160 kg	Observed yield range=0.4 to 8.3 t/ha, R^2=0.90, RMSD=0.8 t/ha. Observed	Explore the relationship between yield

			Bouwing, The Netherland			P(yr)=646 mm. The Bouwing: soil type: silty clay loam, P(yr)=763 mm.	N/ha. PAGV: 80, 140, 180, 240 kg N/ha. The Bouwing: 0, 60, 70, 160, 170, 230 kg N/ha.	biomass range=0.03 to 20 t/ha, R^2=0.97, RMSD=1.2 t/ha. Observed LAI range=0 to 5.5 m2/m2, R^2=0.65, RMSD=1.2 m^2/m^2.	and N-fertiliser application
6	(Fisher et al., 2001)	NWheat	Balla, Wongan Hills, Merredin, East Beverley, Katanning, Newdegate, Esperance, and Salmon Gums, Western Australia, Australia	1989-1992	Spear, Kulin	Wongan Hills: based on 1900-1999 historical weather data, the weather was dominated by summer rainfall<=45 mm, P(t)<=390 mm, early season (April to May) <=140 (n=46) and summer rainfall>45 mm, P(t)>390 mm, early season>140 (n=47).	N treatment of 150 kg N/ha as 90 kg at sowing and 60 kg 4 weeks after sowing. Total of 111 sowing dates between 9th April and 19th July.	Spear: dates of anthesis-RMSD=12.1days, R^2=0.76, bias=-1.0%. Kulin: dates of anthesis-RMSD=9.5days, R^2=0.86, bias=1.4%.	Provide information on choice of cultivar and sowing date
7	(Asseng et al., 2001)	NWheat	Moora, Wongan Hills, Merredin, Western Australia, Australia	Up to 87 continuous years.	Spear, Amery	Sand (PAWC=55 mm), clay soil (PAWC=109 mm). Moora: P(yr)=461 mm, P(gs)=392 mm (mean), 165 to 648 mm (range). Wongan Hills: P(yr)=386 mm, P(gs)=322 mm (mean), 112 to 535 mm (range). Merredin: P(yr)=310 mm, P(gs)=235 mm (mean), 102 to 418 mm (range).	N treatments of 0, 30, 60, 90, 150, 210 kg N/ha. Three sowing dates: DOY 135 (15 May), DOY 155 (4 June), DOY 175 (24 June).	The yield in the Mediterranean climatic region of Western Australia depends on soil water-holding capacity, nitrogen management, rainfall amount and especially, seasonal rainfall distribution.	Explore the water- and nitrogen-use efficiency
8	(Asseng et al., 2002)	NWheat	New South Wales, Australia; Wongan Hills and Cunderdin, Western	1997 (Western Australia)	Spear, Amery	Wongan Hills: P(yr)=391 mm, soil types: sand (PAWC=55 mm). Cunderdin: P(yr)=367 mm, soil type: clay	N treatments of 0, 30, 60, 90 kg N/ha at Wongan Hills experiment.	Western Australia: grain protein-RMSD=1.9 to 2.0%. New South Wales: the model slightly overestimated grain protein which was mainly	Model development and validation

			Australia, Australia			(PAWC=109 mm). New South Wales: P(yr)=436, 536 mm, soil types: loam (PAWC=159 mm).		due to an overestimation in available nitrogen.	
9	(Lilley et al., 2003)	2.1	Harden (HD) and Condobolin (CDBL), NSW, Australia	1989-2000 (HD), 1991-1993 (CDBL)	Janz (HD), Rosella and Dollarbird (CDBL)	HD: Red earth. P(gs)=179 to 539 mm. Tg(yr)=14.7°C. PAWC=145 mm. CDBL: Red brown earth. P(gs)= 234, 341 and 324 mm, Tg(yr)=17.5°C, PAWC=169 mm.	HD: N treatments of 22, 23, 97, 110, 114, 130 kg N/ha) CDBL: N treatments of 10 and 25 kg N/ha	Corresponded well in dynamics other than in 1991, model failed to capture the effects of environmental stresses and overestimate yield and biomass.	Model evaluation
10	(Sadras et al., 2003)*	N/A	Mallee region of Southeastern, Australia, Australia	1998-2002	Hartog	Sandy regolithic, hypocalcic, calcarosols. P(gs)=218 to 351mm. maximum temperature=18 to 19.5°C.	N applied, no nitrogen stress.	When field-measured soil water properties were used, model simulations improved: yield-RMSE from 0.31 to 0.19 t/ha, R^2 from 0.60 to 0.74; SWC-RMSE from $2.1*10^{-3}$ cm^3/cm^3 to $9.3*10^{-4}$ cm^3/cm^3, R^2 from 0.47 to 0.72.	Quantify the effect of environmental factors
11	(Wang et al., 2003)*	N/A	Queensland and Western Australia, Australia	1987-1995	Five local varieties	Soil types: duplex, deep sand, loamy sand.	Four nitrogen treatments, some with water or residue treatments	yield-RMSD=0.74 t/ha, R^2=0.80; biomass-RMSD=1.62 t/ha, R^2=0.82. Could explain >80% of the total biomass/yield/maximum LAI variations.	Model evaluation
12	(Asseng et al., 2004)	NWheat	Obregon, Mexico; Maricopa, USA; Lincoln, New Zealand; Wongan Hills and Cunderdin, Western Australia, Australia	1989-1990 and 1994-1995 (Obregon); 1991 (Lincoln); 1997 (Western Australia); 1992-1994, 1995-1997 (Maricopa)	Yecora70, Batten, Amery, Wilgoyne, Spear	Clay loam, sandy loam, sand, and clay.	Rising temperature (Obregon), increased levels of water deficit (Lincoln), late water deficit (Western Australia), elevated atmospheric CO_2 (Maricopa).	Obregon: yield-RMSD=1.0 t/ha, biomass-RMSD=2.8 t/ha; Lincoln: yield-RMSD=1.2 t/ha, biomass-RMSD=1.9 t/ha, R^2=0.90, LAI-RMSD=1.3 m^2/m^2, R^2=0.53; Western Australia: yield-RMSD=0.5 t/ha, R^2=0.77, biomass-RMSD=1.1 t/ha, R^2=0.86; Maricopa: yield-RMSD=1.1 t/ha, R^2=0.72,	Analyse the climate change impacts on yield

								biomass-RMSD=1.6 t/ha, R^2=0.94, LAI-RMSD=0.9 m^2/m^2, R^2=0.73.	
13	(Yunusa et al., 2004)	1.4 Patch 2	Roseworthy, Minnipa, and Wunkar, South Australia, Australia	1995-1996 (Roseworthy), 1997 (Minnipa, Wunkar)	Janz (Roseworthy, Minnipa, Wunkar), Excalibur (Roseworthy)	Roseworthy: P(yr)=420 mm, P(gs)=320 mm, Tx and Tn=8.2 and 18.8°C. Soil: red brown earth. Minnipa: P(gs)=230 mm, Tx and Tn=7.9 and 19.1°C. Soil: sandy loam topsoil underlayed by calcareous subsoil. Wunkar: P(gs)=170 mm, Tx and Tn=5.5 and 18.8°C. Soil: loamy sand, grading into calcareous sandy clay and heavy clay.	Roseworthy: N treatments of 0, 50, 75, 100 kg N/ha.	Yield- RMSD=0.447 t/ha, R^2=0.69; grain weight RMSD=7.0 mg, R^2=0.31; grain protein RMSD=5.7%, R^2=0.03.	Evaluate yield response to environmental factors
14	(Luo et al., 2005)	2.0	Cummins, Keith, Lameroo, Minnipa, Naracoorte, Orroroo, Roseworthy, and Wanbi, South Australia, Australia	Continuous 100 years	Janz, Excalibur	Cummins, Keith, Naracoorte, and Roseworthy: wetter climates with P(yr)=430 to 580 mm, P(gs)=292 to 397 mm. The other four sites: drier with P(yr)=304 to 388 mm, P(gs)=186 to 251 mm. Cummins: clay loam, PAWC=140 mm. Keith: loamy sand, PAWC=76 mm. Lameroo: fine sandy loam, PAWC=111 mm. Minnipa: sandy loam, PAWC=157 mm. Naracoorte: sandy clay loam,	Fertilised to ensure no nitrogen stress. Changing rainfall, temperature, and CO_2 concentration conditions.	Rainfall is by far the most influential factor on change in median grain yield in the medium to low rainfall areas.	Analyse the climate change impacts on yield

						PAWC=125 mm. Orroroo: sandy loam, PAWC=134 mm. Roseworthy: loam, PAWC=122 mm. Wanbi: sandy loam, PAWC=132 mm.			
15	(Paydar et al., 2005)*	N/A	Northern NSW, Australia	1995-1998	N/A	Black vertosol, P(yr)=684mm. The available moisture holding capacity of the soil is large (505 mm to 3 m depth).	N treatment of 100 kg N/ha.	The yield predictions are generally good for wheat, barley, and sorghum but less so for legumes. Failed to predict rapid increases in subsoil moisture.	Quantify the effects of different cropping systems on the water balance
16	(Oliver et al., 2006)*	N/A	Buntine, northern sandplain of the Western Australia wheatbelt	1997, 1999, 2002-2005	N/A	P(yr)<400mm, PAWC=32 to 110mm, 50 kg N/ha in the soil profile.	N applied from 0 to 150 kg N/ha at sowing.	Yield-RMSD=0.518 t/ha.	Explore the importance of PAWC as a driver of yield variation
17	(Hunt et al., 2006)*	Yield Prophet	338 paddocks of 236 growers in Australia	1997, 1999, 2002-2005	N/A	N/A	N/A	Paddocks with appropriate measured soil characterisation and soil profile samples: R^2=0.68, 68% of simulated results were within 0.5 t/ha. Paddocks without appropriate soil information: R^2=0.54, 49% of simulated results were within 0.5 t/ha.	Explore the importance of PAWC as a driver of yield variation
18	(Wong and Asseng, 2006)	NWheat	Three Springs, Western Australia	1998-2002	Blade, Brookton, Carnamah	P(yr)=445 mm, of which 370 mm falls in growing season (May to October).	N treatments of 0, 60, 150 and 210 kg N/ha	Yield-RMSE=1.0 t/ha.	Develop an method to use APSIM spatially
19	(Moeller et al., 2007)	4.2	Dry areas at Tel Hadya, northwestern Syria	1998-2000	Cham3	Semi-arid, continental Mediterranean climate with cool, wet winters and hot, dry summers. P(yr)=340mm, Tg(yr)=17.6 °C.	N treatments of 0, 60 and 100 kg N/ha. Irrigation: 0 and 342 mm.	At pre-anthesis stage, the model overestimated leaf-area, nitrogen uptake and biomass accumulation.	Model parameterisation and validation

						Growing season: early/mid November to early/late May. Vertisol and Inceptisol Soil group. pH is around 8. The soil organic matter content is mostly lower than 1% in the 0 to 0.20 m layer.			
20	(Hochman et al., 2007)	5.0	Southern Queensland, northern NSW	2003-2004	Baxter, H45, Wollaroi, Yallaroi, Babbler, Hybrid Meteor, Strzelecki, Sunbrook	PAWC=93 to 120 mm	No significant weeds, pests, diseases, or nutrient deficiencies experienced.	When measured PAWC data were used, yield-RMSD=0.5 t/ha, R^2=0.82. When calculated instead of measured PAWC were used, yield-RMSD=0.78 t/ha, R^2=0.69. When calculated PAWC were used and kl was adjusted, yield-RMSD=0.53, R^2=0.84.	Evaluate yield response to environmental factors
21	(Lilley and Kirkegaard, 2008)*	5.0	Gundibindyal, NSW, Australia	2000-2004	Janz	In-crop rainfall=440 mm, PAWC=173 mm.	N treatment of 178 kg N/ha.	Yield-RMSD=0.4 t/ha, R^2=0.90.	Evaluate yield response to environmental factors
22	(Lawes et al., 2009)	5.2	Buntine, Western Australia	1997-2005	Calingiri, Brown manured, Westonia, Wyalkatchem	P(yr)=300 to 400 mm. PAWC=52 to 131 mm.	N applied from 20 to 60 kg N/ha. Sowing dates ranged from 15th May to 2nd June.	Yield-RMSE=0.31 t/ha, R^2=0.86.	Explore the yield-PAWC relationship
23	(Oliver et al., 2009) (assembled datasets)	5.2	Wheatbelt of Western Australia	1996-2006	Various varieties	P(yr)=300 to 500 mm, P(gs)=243 mm ± 88 mm. PAWC=33 to 434 mm.	No nitrogen stress	Yield-RMSE=0.455 t/ha, R^2=0.78. Model slightly overpredicted yields.	As a predictive tool benchmark
24	(Wang et al., 2009)*	5.3	Luancheng, Yucheng and Fengqiu, NCP, China	1997-2006	Gaoyou503, Zhixuan1, Keyu13, Zhengmai9023	Luancheng: P(yr)=481 mm. PAWC=335 mm. Yucheng: PAWC=341 mm. Fengqiu: PAWC=204 mm.	Luancheng: Irrigated of 202 to 404 mm. Yucheng: N treatments: 182 to 215 kg N/ha, irrigated of 110,	Yield-RMSD=0.80 t/ha, R^2=0.66.	Evaluate yield response to environmental factors

							152, 230, 400 mm.		
25	(Anwar et al., 2009)	5.3	Victoria, Australia	2004-2006	Yitpi	P(yr)=354 mm, P(gs) (April-October) =239 mm. The texture of surface soil varying from loamy sand to sandy clay loam and subsoil varying from sandy loam to sandy clay.	N treatments of 0, 14, 26, 50 kg N/ha (2004); 0, 25, 33, 50 kg N/ha (2005).	Yield-RMSE=0.31 t/ha, R^2=0.96.	Evaluate yield response to environmental factors
26	(Bell et al., 2009)* (assembled datasets)	5.4	Western Australia, Queensland and New Zealand	2002-2005, 1980s-1990s	Wyalkatchem	P(yr)=300 to 595 mm. Red loam, deep sand and shallow gravel. PAWC=0 to 75 mm.	N treatment of 100 kg N/ha.	Yield-RMSD=0.537 t/ha, biomass-RMSD=1.27 t/ha, which representing 18% and 17% of the mean observed values.	Compare the profits of yield harvesting and sacrificing the crop to grazing
27	(Oliver and Robertson, 2009)	N/A	Wheatbelt of Western Australia	2004-2006	Wyalkatchem, Carnamah, Bonnie Rock, Westonia, Yitpi, Calingiri	P(yr)=308 to 446 mm (P(gs)is about 75% to 86% of P(yr)). PAWC=59 to 208mm.	N applied from 10 to 74 kg N/ha.	Yield-RMSE=0.538 t/ha, R^2=0.80.	Evaluate yield response to environmental factors
28	(Hochman et al., 2009a) (assembled datasets)	Yield Prophet (based on APSIM 6.1)	Australia (344 winter wheat crops. Victoria (176), South Australia (75), New South Wales (43), Western Australia (38) and Queensland (4))	2004-2007	Various varieties	A range of soil types, from shallow sands (PAWC=22 mm) to deep Vertosols (PAWC=279 mm).	N applied from 32 to 588 kg N/ha (mean=124 kg N/ha). Irrigated.	Yield-RMSD=0.8 t/ha, R^2=0.71.	Model evaluation based on a large dataset from on-farm crops
29	(Carberry et al., 2009) (assemble datasets)	Yield Prophet	Over 700 commercial crops (wheat=495) in Australia	1992-2007	Various varieties	Different sowing dates, soil types	N/A	Yield-RMSD=0.19 to 0.80 t/ha, R^2=0.52 to 0.89	Model evaluation
30	(Chen et al., 2010a)*	5.1	NCP, China	2000-2001	Gaoyou503	Loam soil, with texture ranging from sandy loam in surface layers to	Urea applied of 150 kg/ha. Irrigated from 80 to 118 mm.	The model was able to explain more than 90% yield variation.	Explore optimal water management strategies

						light/median loam at 40 to 80 cm depth and to light clay below 80 cm.			
31	(Chen et al., 2010b)*	5.3	NCP, China	1997-2002, 2004-2006	Gaoyou503, Zhixuan1, Keyu13, Zhengmai9023	P(yr)=481 to 615 mm	N applied to ensure no nitrogen stress. Irrigated from 0 to 420 mm at critical growing stages.	Yield-RMSE=0.83 t/ha, biomass-RMSE=1.40 t/ha.	Evaluate yield response to environmental factors
32	(Chen et al., 2010c)*	5.3	NCP, China	1997-2006	Zhixuan1	P(yr)=481 to 615 mm	N applied and irrigated (lacked accurate irrigation and fertilisation records).	No significant systematic over- or under-estimations was found when predicting LAI, biomass and yield. LAI-d index=0.85, biomass-d=0.92, yield-d=0.96. LAI-R^2=0.61, biomass-R^2=0.62, yield-R^2=0.88.	Evaluate yield response to environmental factors
33	(Holzworth et al., 2014) (assembled datasets)	N/A	Assembled datasets	N/A	Various varieties	Range of soil types, locations, sowing dates.	N/A	Yield-RMSE=0.46 t/ha, R^2=0.93.	Model development and validation
34	(Balwinder-Singh et al., 2011)*	5.1	Punjab, India	2006-2008	PBW343	Clay loam soil. P(gs)=88, 159 mm.	N applied to ensure no nitrogen stress. With and without mulch. Six irrigation scheduling treatments, including 75, 150, 225 mm.	Mulch: yield-RMSE=0.443 t/ha, NRMSE=12.4%, R^2=0.91. Biomass-RMSE=0.3 t/ha, NRMSE=3.6%, R^2=0.99. Non-mulch: yield-RMSE=0.55 t/ha, NRMSE=16.5%, R^2=0.86. Biomass-RMSE=0.8 t/ha, NRMSE=10.8%, R^2=0.92.	Model calibration and evaluation
35	(Lobell et al., 2012)	N/A	Indo-Gangetic Plains, India	2000-2009	Zippy	March and April have averaged 20 to 30 days of daily temperature exceeded 34°C (grain filling period of wheat).	Average N treatment of 145 kg N/ha. Irrigated.	APSIM underestimated potential yield losses for warming in the study area.	Evaluate yield response to environmental factors

36	(Mohanty et al., 2012)	6	Bhopal, India	2002-2006	Sujata	P(t) of 2002: 763 mm; 2003: 1113 mm; 2004: 863 mm; 2005: 917 mm	Three N treatments: (1) no nutrient added, (2) 100 kg/ha N, 22 kg/ha P, 17 kg/ha K, (3) 16 t/ha farmyard manure. Irrigated of 80 or 240 mm.	SWC-R^2=0.71 to 0.88. The model realistically predicted yield and nitrogen uptake.	Model calibration and evaluation
37	(Zhang et al., 2012)	6.1	Shangzhuang, Quzhou, Huangfanqu, at NCP, China	2009-2010	Nongdan211, Han6172, Yanzhan4110	Shangzhuang: P(yr)=104, 114 mm, Tg(yr)=8.0, 5.9°C. Quzhou: P(yr)=260, 128 mm, Tg(yr)=9.8, 7.7°C. Huangfanqu: P(yr)=307, 313 mm, Tg(yr)=10.4, 9.6°C.	N applied of 240 kg N/ha and irrigated of 120 mm. Delayed sowing dates or decreased planting density.	Increased error of simulated yield with (1) delayed sowing dates: yield-NRMSE=7 to 12% (0.29 to 0.57 t/ha), 11 to 16% (0.65 to 1.09 t/ha), 16 to 22% (0.56 to 0.97 t/ha); (2) decreased planting density: yield-NRMSE=9 to 12% (0.54 to 0.56 t/ha), 11 to 12% (0.72 to 0.90 t/ha), 16 to 19% (0.77 to 1.26 t/ha).	Model evaluation
38	(Hochman et al., 2013)	N/A	Wimmera, Victoria, Australia	Continuous 26 years	Yitpi	45 stations with 5 soil types.	N of 50 kg N/ha was applied whenever soil nitrate in the root zone falls below 50 kg N/ha.	Annually estimated yield gaps of 0.66 to 4.12 t/ha with an average yield gap of 2.0 t/ha.	Quantify yield gaps
39	(Wang et al., 2013)	5.3	NCP, China	1980-2009	Multiple local varieties	Tg(yr)at six sites were 13.0, 12.9, 13.0, 14.1, 14.5 and 15.1°C, while P(t) were 515, 538, 535, 588, 550 and 995 mm, respectively.	Flood irrigated three to four times of 250 to 300 mm.	The model slightly overestimated the days to jointing and flowering and underestimated the days to emergence and maturity dates.	Explore the phenological trends
40	(Zhang et al., 2013)*	6.1	NCP, China	2009-2010	Nongda211, Han6172, Yanzhan4110	Mean minimum temperature of −8·8 to −3.8°C in January.	N applied of 240 kg N/ha and irrigated of 120 mm.	Underestimate yield of 0.4 to 0.6 t/ha, RMSE=0.5 to 0.9 t/ha.	Evaluate the climate change impacts on yield
41	(Carberry et al., 2013) (assembled datasets)	Yield Prophet	849 commercial wheat crops in Australia	2004-2011	Various varieties	P(yr)=182 mm	N applied. Rainfed.	APSIM was able to closely simulate commercial wheat yield.	Quantify yield gaps

42	(Brown et al., 2014) (assembled datasets)	N/A	28 cropping sites in Australia, USA, New Zealand, and NCP, China	N/A	Various varieties	P(yr)=227 to 839 mm	N treatments appied ranged from 0 to 325 kg N/ha. Irrigated or rainfed.	Biomass-R^2=0.93, grain yield-R^2=0.92, biomass nitrogen-R^2=0.87, grain nitrogen- R^2=0.87.	Model development and validation
43	(Bryan et al., 2014)*	7.3	Northern NSW region, Western Australia, South Australia Victoria region	2006	N/A	N/A	N applied of 225 kg N/ha.	Census-reported yield=0.54 to 2.31 t/ha (median=1.26 t/ha), simulated yield=0.639 to 2.906 t/ha (median=1.553 t/ha).	Evaluate yield response to environmental factors
44	(Peake et al., 2014)*	7.4	Queensland, Australia	2008-2009	EGA Gregory, Kennedy, Ventura, Strezelecki, Baxter	Crops experienced lodging, water stress, high temperature, hail damage in 2008. In 2009, lodging, water stress, moderate to severe nitrogen stress were also observed.	2008 crops were fertilised. All crops were irrigated.	APSIM accounted for 72% of the non-lodged wheat yield variation and a RMSD=1.08 t/ha. While overestimated lodged crop yield and underestimated crop in the low-nitrogen field.	Quantify yield gaps and yield response to environmental factors
45	(He et al., 2014)	7.4	Loess Plateau, China	2007-2008	Changwu89134	Average Tx=15.2 to 17.1°C. Average Tn=2.4 - 6.4°C. P(yr)=320.8 – 479.8 mm.	Urea applied of 300 kg/ha. Rainfed.	LAI-d index=0.91, R^2=0.89. Biomass-d=0.96, R^2=0.91. ESW-d=0.94, R^2=0.78. ET-d=0.95, R^2=0.85.	Evaluate yield response to environmental factors
46	(Wang et al., 2014)*	7.5	Northern China	1989-2003	N/A	P(t)=262, 608, 630, 848 mm for each site.	Three N application scenarios: (1) fertilised and irrigated, (2) stubble managed, fertilised and irrigated, (3) control (rainfed and irrigated, without fertiliser and stubble).	Reasonably simulate 50% to 90% of the yield variation.	Evaluate yield response to environmental factors

47	(Zhao et al., 2014a)	7.5	NCP, China	2009-2011	SJZ15	1961-2010: P(yr)=550 mm, Tg(yr)=12.9°C.	N application ranged from 0 to 330 kg N/ha. Irrigation ranged from 75 mm to 375 mm.	The model overestimated biomass and yield. The calibrated nitrogen concentration improved biomass and nitrogen updated simulations, especially under low nitrogen input.	Evaluate the threshold nitrogen concentration used in the model
48	(Zhao et al., 2014b)*	7.5	NCP, China	2003-2011	SJZ8, SJZ15	Summer monsoon climate, with P(yr)=550 mm, Tg(yr)=12.9°C (1961-2010). Calcaric Fluvisol with a sandy clay loam texture.	N treatments of 0, 123, 158, 192, 261, 330 kg N/ha. Irrigation: 3 × 75=225 mm.	The model accounted for more than 85% of the biomass variation, biomass-RMSE=1.1 t/ha, more than 80% of the yield variation, yield-RMSE=0.73 t/ha.	Evaluate the root modelling
49	(Xiao and Tao, 2014)*	N/A	Northern China	2005-2009	Fengkang7, Jingdong8, Hengshui741, Shimai12, Boai7422, Zhengmai9023, Fu63, Lumai23	Tg(yr)from 12.8 to 15.7 for all four locations.	N applied: 90 or 120 kg N/ha used as base fertiliser and 60 or 75 kg N/ha added at jointing stage. Irrigated 4 × 50 mm.	The average difference between modelled and observed yield<0.5 t/ha, R^2=0.85.	Evaluate yield response to environmental factors
50	(Li et al., 2014)	N/A	NCP, China	1981-2010	Jinfeng1, Jiamai26, Gaoyou503, Bainong3217, Yumai18, Zhengmai9023, Jinan13, Lumai15, Lumai21	N/A	Local traditional practices: irrigation was not conducted every year, but fertiliser was used several times every year.	Yield-RMSE=0.3205 to 0.8291 t/ha, NRMSE (%) =5.7 to 14.1, d-value=0.90-0.97, R^2=0.71-0.89.	Identify the change pattern of yields
51	(Soltani and Sinclair, 2015) (assembled datasets)	7.X	Grogan, Iran	2005-2008	Several local varieties	Silty clay. P(gs)=340 mm, average Tx=17.2°C, average Tn=7.3°C.	N applied from 0 to 122 kg N/ha. Part of crops were irrigated. 8 to 12 sowing dates. 6 to 7 sowing densities ranged from 50-800 plant/m^2.	LAI at anthesis: RMSE=0.74, r=0.53; dry mass at anthesis: RMSE=1.50 t/ha, r=0.51; dry mass at maturity: RMSE=2.44 t/ha, r=0.72; yield: RMSE=0.62 t/ha, r=0.81.	Model intercomparison and evaluation

52	(Sun et al., 2015)	7.0	NCP, China	2006-2012 (1st experiment), 1984-2012 (2nd experiment)	Jimai7, Jimai36, Jimai733	The total evapotranspiration during the growing season is 400 to 450 mm. Soil: loam, the average water holding capacity is 38%, and the wilting point is 13%.	N application: 100 to 125 kg N/ha (1984 to 1990); 220 kg N/ha (in the 1990s); 250 kg N/ha (after 2000). Irrigation: 1st experiment: full irrigation, critical stage irrigation, minimum irrigation, rainfed. 2nd experiment: the irrigation was managed similarly as the full irrigation treatment.	1st experiment: the model was able to explain more than 83% of the yield variation. RMSE values under full irrigation, critical stage irrigation, minimum irrigation and rainfed were 0.330, 0.567, 0.923 and 0.762 t/ha. 2nd experiment: yield-RMSE=0.590 t/ha.	Evaluate yield response to environmental factors
53	(Acuña et al., 2015)	7.1	10 sites, Tasmania, Australia	1980s, 2020s	Brennan, Isis, Machellar, Revenue, Tennant (winter wheat) and Kellalac (spring wheat)	P(yr)=499 to 965 mm, Tx=16.1 to 17.6˚C, Tn=4.6 to 8.2˚C.	N application ranged from 24 to 245 kg N/ha.	Yield-RMSE=1 t/ha, R^2=0.84.	Explore the potential management strategies to close the yield gap
54	(Deihimfard et al., 2015)*	7.2	Northeastern Iran	2009-2011	Late maturing: Sionz, Gascozhen; early maturing: Chamran	P(gs)=137 to 298 mm, Tx=9.7 to 14˚C, Tn=6.9 to 10.1˚C, Tg(yr)=8.8 to 12.3˚C.	N applied at four levels: 0, 55, 110, 172 kg N/ha. Irrigated 5 to 9 times with 50 mm each time.	Yield-RMSE=0.71 t/ha, R^2=0.83.	Quantify yield gaps
55	(O'Leary et al., 2015)	7.4	Victoria, Australia	2007-2009	Yitpi	Elevated CO_2 condition (550 µmol/mol) and normal CO_2 condition (365 µmol/mol)	N applied of 0 and 53 to 138 kg N/ha. Irrigated/rainfed. Normal and late sowing dates.	APSIM tended to overestimate LAI at DC65 (R^2=0.24, RMSE=0.70 m^2/m^2), biomass at DC31 (R^2=0, RMSE=1.592 t/ha), biomass at DC65 (R^2=0.56, RMSE=1.542 t/ha) and yield (R^2=0.20, RMSE=1.294 t/ha).	Evaluate the climate change impacts on yield
56	(Innes et al., 2015)	7.5	Australia	1982-2008	Hartog	P(gs)=250 to 400 mm. Recurrent	N applied of 69 kg N/ha.	Yield variation (%): RMSE=18.9%, R^2=0.69.	Evaluate model under high-

						drought, high temperature and low rainfall.			temperature episodes
57	(Zhao et al., 2015)*	7.5	NCP, China	2009-2010	SJZ8, SJZ15	Summer monsoon climate, with P(yr)=550 mm, Tg(yr)=12.9°C (1961-2010). Calcaric Fluvisol with a sandy clay loam texture.	1st experiment: 3 × 75 mm irrigations. 0, 123, 192, 261, 330 kg N/ha N applied. 2nd experiment: 1/2/3/5 × 75 mm irrigations. 158 kg N/ha N applied.	Yield-RMSE=0.33 t/ha, R^2=0.97. The simulated grain yield remained similar with modified parameters.	Analyse the resource use efficiency
58	(Ahmed et al., 2016)	N/A	Islamabad, Pakistan	2009-2011	Tatara, NARC-2009, Sehar-2006, SKD-1, F-Sarhad	High rainfall (P(yr)> 1000 mm). Tg(yr)=21.3°C. The annual potential evapotranspiration is about 1600 mm.	N/A	Phenology-RMSE=2.03 to 5.09day, R^2=0.8, maximum LAI-RMSE=0.14 to 0.32 m^2/m^2, R^2=0.83, accumulated biomass-RMSE=0.15 to 0.40 t/ha, R^2=0.92 and yield-RMSE=0.12 to 0.31 t/ha, R^2=0.82.	Model calibration and evaluation
59	(Van Oort et al., 2016)	7.4	NCP, China	2006-2007	Shimai12	Continental monsoon: cold and dry winters. P(yr)=533 mm, only 2% occurs in winter.	Three levels of irrigation: (1) 0, (2) 75 mm water at stem extension, (3) 75 mm at stem extension plus 75 mm water at booting. Enough N was applied to ensure no nutrient limitation.	The model was able to explain 95% of biomass variation, 90% of LAI variation, 84% of SWC variation, 82% of yield variation. Biomass-RMSE=0.88 t/ha, LAI-RMSE=0.72 t/ha, SWC-RMSE=27 mm, yield-RMSE=0.64 t/ha.	Construct groundwater neutral cropping systems
60	(Li et al., 2016)	7.5	NCP, China	2008-2010	Jimai22	Typical temperate monsoon climate. Tg(yr)=13.9 °C, P(yr)=547 mm for the period 1990-2010.	Four N treatments: (1) 0, (2) farmer conventional fertilisation (234 kg/ha urea applied), (3) reduced	APSIM explained 94% and 88% of the variation in final biomass and grain yield, with RMSE of 1.28 and 0.82 t/ha.	Explore possibility of resources usage reduction while maintaining the yield

							fertilisation (144 kg/ha urea applied), (4) reduced N with manure (54 kg/ha chicken manure + 90 kg/ha urea applied).		
61	(Mielenz et al., 2016)*	7.5	Kingaroy (KGR) and Kingsthorpe (KTHP), Southeastern Queensland, Australia	2009 (KTHP), 2011(KGR)	Hartog (KGR), Lang (KTHP)	Humid subtropical. Tg(yr)=18.2 °C at both sites. P(yr)=776 mm and 630 mm, respectively.	KGR: sprinkler irrigated, four N treatments were applied: (1) 0, (2) 20 kg/ha of urea applied adjusted according to estimated residual soil N, (3) 80 kg/ha of urea applied, (4) 140 kg/ha of urea applied. KTHP: fertilised, three irrigation treatments: when 50/60/85% of the PAWC was depleted.	R^2=0.92.	Identify strategies for mitigating crop N2O emissions
62	(Zeleke and Nendel, 2016)*	7.6	NSW, Australia	2013-2014	EGA Gregory, Livingston	Sandy clay loam Red Kandosol. For 2013 and 2014: number of frosts=40 days, 48 days; P(t)=263 mm, 326 mm.	Irrigation applied=247 mm (in 2013), 229 mm (in 2014).	Yield-RMSE=0.65 t/ha, R^2=0.92.	Evaluate yield response to environmental factors
63	(O'Leary et al., 2016)*	7.6	Wagga Wagga, New South Wales and Warwick, Queensland, Australia	1979-2003,1968-2012	N/A	In some seasons, crops suffered from diseases like root-lesion nematodes.	Different stubble, tillage and nitrogen application managements applied.	Wagga Wagga: yield-RMSE=1.08 t/ha. Warwick: yield-RMSE=1.39 t/ha.	Evaluate yield response to environmental factors
64	(Liu et al., 2016a)*	7.7	Environment-controlled	2010-2014	Yangmai16, Xumai30	Environment-controlled phytotron experimental	N and irrigation were applied to	Heat happened at anthesis: yield-R^2=0.73. Heat	Evaluate the model ability of

			chamber in China			datasets under heat stress at anthesis and grain filling stages. The plant density=10 plants per pot, the diameter of a pot=0.28m.	ensure no water or nitrogen stress.	happened at grain filling: yield-R^2=0.46.	simulating heat impacts
65	(Araya et al., 2017)	7.4	Ethiopia	2011-2012	HAR-2501	N/A	Two levels of N applied: (1) 0, (2) 64 kg N/ha. Rainfed.	Simulated phenology R^2 over 0.8, 6.0day, yield-R^2=0.63, 0.14 t/ha.	Evaluate the climate change impacts on yield
66	(Gaydon et al., 2017)* (assembled datasets)	N/A	Twelve Asian countries, total of 43 experimental datasets, 966 crops (326 wheat)	Various years	Various varieties	Different weather conditions (temperature, rainfall, CO_2 level).	Different sowing dates, dates of transplanting, N and surface residue treatments, rainfed or irrigation conditions.	Yield-RMSE=0.845 t/ha, R^2=0.79, standard deviation=1.794 t/ha. APSIM underestimated LAI, biomass, and yield in NCP, China due to incorrect temperature response of physiological processes.	Model evaluation
67	(Zhao et al., 2017)*	N/A	Inner Mongolia, China	2011-2014	Spring wheat variety	Tg(yr)= -1 to 10˚C, P(yr)=50 to 450 mm.	N/A	Yield-RMSE=0.029 t/ha to 0.208 kg/ha, NRMSE=0.92%-6.4%, d-index=0.85 to 0.95.	Evaluate the climate change impacts on yield
68	(Holzworth et al., 2018)	Next Generation	Various locations	Wheat model: 650 simulation years	Various varieties	N/A	N/A	Included in the model files.	Model development and validation
69	(Hussain et al., 2018)	7.8	Faisalabad and Layyah in Punjab-Pakistan	2013-2015	Lasani-2008, Punjab-2011, Galaxy-2013	T(gs)= -0.1 to 43˚C	Eleven planting dates (16th October to 16th March with interval of 15 to 16 days). Irrigation was applied to ensure no water stress. N applied of 120 kg N/ha.	The model overestimated yield with late planting dates.	Model intercomparison and evaluation
70	(Phelan et al., 2018)*	7.8	Tasmania, Australia	2005-2010	Mackellar_Tas, Revenue, Tennant	Cressy, Epping forest, Symmons plains: P(yr)=628 mm, Tx=17.2,	N applied of 75 kg N/ha. Rainfed.	Yield-R^2=0.83, MPE (mean prediction error) =11%, EF=0.82, v (ratio of	Produce data for further incorporation

						Tn=5.1. Soil: fine sandy loam (PAWC=217 mm), clay loam (96 mm), loam (PAWC=221 mm).		variance in measured to simulated values) =1.09.	into another model
71	(Brown et al., 2018) (assembled datasets)	7.9	Eight countries	N/A	Various varieties	48 experiments, 655 treatments, different planting years.	Different time of sowing, N fertilizer, irrigation, residue additions, population, tillage.	$R^2 \geq 0.84$ and NSE ≥ 0.81 for all model variables presented (R^2=0.84, RMSE=1.005 t/ha and NSE = 0.81 for yield), except grain protein which had an R^2 of 0.42 and a NSE of 0.36. Flag leaf-R^2=0.98, anthesis-R^2=0.98.	Model development and validation
72	(Bahri et al., 2019)	N/A	Nabeul, Cherfech, Hendi Zitoun, Boulifa, Oued Mliz, and Mornag, Tunisia	1989-1992, 1996-1998, 1999-2000, 2003-2006	Karim	Nabeul: P(gs)=232 mm, soil: sandy; Cherfech: P(gs)=345, 516 mm, soil: silty clay loam; Hendi Zitoun: P(gs)=299, 139, 105 mm, soil: silty clay; Boulifa: P(gs)=424, 499, 520 mm, soil: silt-clay sandy; Oued Mliz: P(gs)=240 mm, soil: clay loam; Mornag: P(gs)=125 mm, soil: clay loam.	Nabeul: irrigation=228 mm, N application=132 kg N/ha; Cherfech: irrigation=255, 163 mm, N application =198, 132 kg N/ha; Hendi Zitoun: irrigation=290, 300, 250 mm, N application =60, 150 kg N/ha; Boulifa: irrigation=0, N application =76, 150 kg N/ha; Oued Mliz: irrigation=100 mm, N application =150 kg N/ha; Mornag: irrigation=252 mm, N	Yield-RMSE=1.647 t/ha, agreement index=0.83.	Evaluate yield response to environmental factors

							application =150 kg N/ha.		
73	(Bai et al., 2020)	7.7	NCP, China	1981-2015	Jimai22, Jining142, Zhengzhou761	P(gs)=100 to 300 mm.	On-farm: N fertiliser were applied at sowing and jointing. Three to four times irrigation. High yield: Four irrigation scenarios: (1) no irrigation, (2) one irrigation, (3) two irrigations, (4) three irrigations. Two N applications: (1) one N application at sowing stage of 0 to 300 kg N/ha, (2) Split N application (sowing and jointing stages) of 0 to 300 kg N/ha.	Simulation of observed high yield records: RMSE=1.15 t/ha, NRMSE=12%. Simulation of on-farm yield: RMSE=0.576 t/ha, NRMSE=8.8%.	Quantify yield gaps and seek for options to increase yield
74	(Araya et al., 2020)*	7.7	Kulumsa, Oromia (KARC); Hagereselam, Tigray region (HS); Ilala (IL); Wukro, Tigray region (WU), Ethiopia	2006-2008, 2012 (KARC), 2014 (HS, IL, WU)	Early, medium and late maturing varieties	Four sites: black vertisol, clay soils. P(yr)=820, 669, 583, 565.6 mm. P(gs)=503, 542, 491, 335.4 mm. Average Tx=23.1, 22.5, 23.6, 28.0˚C. Average Tn=10, 11.1, 12.1, 11.1 ˚C.	Two N application rates: (1) 64 kg N/ha, (2) 128 kg N/ha. Rainfed.	Yield-NRMSE (normalised RMSE) =22.8%, days of flowering-RMSE=4.3%, days of maturity-NRMSE=8.3%.	Evaluate the climate change impacts on yield
75	(Yan et al., 2020)*	7.9	NCP, China	2007-2016	KN199	Loamy soil. P(gs)=50 to 230 mm, Tg(yr)=12.7˚C.	Three irrigation treatments: (1) full irrigation (225 to 375 mm), (2) critical stage irrigation (75 mm at jointing stage in addition to	The model could explain approximately 90% of phenology, biomass accumulation, grain yield and seasonal evapotranspiration for winter wheat. The yield-RMSE were 0.263, 0.598	Explore possibility of resources usage reduction while maintaining the yield

							minimum irrigation), (3) minimum irrigation (keep the top 50 cm soil layer above 75% of field capacity).	and 0.453 t/ha under the minimum irrigation, critical stage irrigation and full irrigation treatments.	
76	(Fletcher et al., 2020)*	7.8	Western Australia	2010, 2015	N/A	Rainfed (water limited condition)	N/A	Yield-RMSE=0.77 t/ha, R^2=0.69.	The climate change impacts on the distribution of Australian wheat belt

Overall, researchers report that site-specifically calibrated APSIM-Wheat provides a useful yield prediction tool for a wide range of environments. Nevertheless, while the model incorporates stress functions to account for limitations of water, nitrogen, heat and frost (Zheng et al., 2014), it sometimes fails to capture these stress effects sufficiently (Barlow et al., 2015). Each of the stress effects is now discussed in more detail.

2.5. Factors affecting APSIM yield prediction

Several factors affecting APSIM-Wheat yield prediction were distilled and presented in the following section after we reviewed papers in Table *3*. Identified influencing factors include model calibration, crop resources (water, nutrition), temperature and biotic stress.

2.5.1. Model calibration

APSIM-Wheat performs optimally when reliable and accurate soil information is available and biotic/abiotic stresses are absent (Dalgliesh et al., 2012). Accurate specification of soil water holding characteristics affects APSIM-Wheat prediction performance (Lilley et al., 2003; Sadras et al., 2003). Specifying lower limits of plant available water with field measurements rather than using estimations from soil texture can improve simulation accuracy. In one study, the R^2 of the relationship between simulated and observed yields increased from 0.60 to 0.74, and the RMSE decreased from 0.31 t/ha to 0.19 t/ha when using lower limits of extractable water derived from field gravimetric soil water measurements , compared with texture based estimates (Sadras et al., 2003). Hunt et al. (2006) indicated that when the model was initialised with appropriate soil water holding characteristics and input data, 68% of the yield predictions were within ±0.5 t/ha of the observed yields.

Across the Australian dryland cropping area, the crucial challenge for predicting commercial wheat yield is to accurately describe the soil characteristics, soil water and nitrogen sources (Carberry et al., 2009). This requirement motivated the development of the APSoil soil database (Dalgliesh et al., 2012, 2009), which provides representative soil parameters for major Australian soils. For Australian paddocks, if field measured soil parameters are not available, APSoil can provide soil information such as the Plant Available Water Capacity (PAWC) based on approximate soil type information (Innes et al., 2015; Phelan et al., 2018).

Some other parameters and functions in APSIM have been modified by authors to achieve better performance. The maximum and critical nitrogen concentration in leaves used in the APSIM-Wheat model was too low when compared to the observed data collected from NCP fields. Adjustment of these two parameters can improve the model simulation, especially under low nitrogen input (Zhao et al., 2014a). Root growth parameters were modified to better simulate the root biomass and its distribution (Zhao et al., 2015, 2014b). The soil moisture factor used for the denitrification rate calculation was modified by Mielenz et al. (2016), instead of using drained upper limit (DUL) as the threshold, the authors modified it to be decided by the water-filled pore space and saturation (SAT). Brown et al. (2018) pointed out that the phenology model needs careful parametrisation for different cultivars.

2.5.2. Water stress

Balwinder-Singh (2011) evaluated APSIM-Wheat in India for different management approaches with and without mulch and six irrigation scheduling treatments. The results indicated that the model underpredicted grain yield by 0.6–1 t/ha when crops were subject to water deficit. Asseng et al. (1998b) attributed the underpredicted yield to insufficient re-translocation of stored pre-anthesis carbohydrates to the grain by APSIM. They suggested the model can be improved by including functions to remobilise additional carbohydrates of stem into the grain when crops experience severe drought conditions. In a separate study, water-limited simulation resulted in overestimation of yield by 0.55 t/ ha with a mean simulated yield of 2.3 t/ha in the Western Australian wheat belt (Fletcher et al., 2020).

2.5.3. Heat stress

Heat stress during wheat growth, especially at anthesis and grain filling stages, affects APSIM yield prediction significantly (Liu et al., 2016a). Hochman et al. (2009) reported that a widespread unseasonal heatwave, followed by a frost in the Wimmera and Mallee regions of Victoria, Australia in 2004 caused the model to overestimate yields by 0.9 t/ha with a mean simulated yield of 1.8 t/ha. Lobell et al. (2012) found that the shortening of the green season (by +2°C warming) was underrated by APSIM by up to eight days, and yield losses were underestimated by up to 50% after comparing the model simulation with a regression model based on nine years of wheat phenology (from satellite observations) and daily temperature data. Liu et al. (2016a) conducted environment-controlled chamber experiments to test the model response when heat stress happened at anthesis and grain filling stages. The results indicated that wheat is more sensitive to heat at anthesis since both grain number and size are affected, while heat during grain filling only decreased the grain size, due to a shorter grain filling duration. The model failed to capture the heat stress impacts at anthesis and grain filling on grain number, underestimating the impacts on grain size at both stages. Hussain et al. (2018) evaluated performance of APSIM simulations of winter wheat sown at different times, from early to extremely late. The model poorly predicted yield for late planting dates due to high temperature during grain filling. Even a short-term exposure of wheat to extreme high temperatures at early grain filling can reduce the duration of grain filling and hence the cumulative degree days and resulting in smaller harvest yield (Stone and Nicolas, 1995). Lobell et al. (2012) also detected greater senescence from extreme heat, beyond the impacts of increased average temperatures.

In summary, the quality of grain number and size simulation exerts a critical influence on the accuracy of yield prediction. Only using the daily mean temperature to apply heat stress is not effective in accounting for heat wave impacts. In addition, short periods (1–3 days) of extremely high temperatures (> 33˚C) can also affect the crop growth and ultimately result in a significant reduction in grain yield (Barlow et al., 2015). Accounting for high daily maximum temperatures as another variable to determine the heat stress impact would help the model better respond to heat waves.

2.5.4.Frost damage

Barlow et al. (2015) summarised three crucial physiological damages that have impacts on yield production in response to a frost event: seedling death during the vegetative stage, sterility at anthesis and death of formed grains during grain filling. Frost during the vegetative stage has smaller impact on harvest yield than during later stages as it mainly affects seedling survival (Fuller et al., 2007) and causes leaf senescence (Shroyer et al., 1995). The greatest yield production impacts resulting from frost are at the reproductive stage, and this frost sensitivity increases from heading to the end of anthesis (Frederiks et al., 2012).

Hochman et al. (2013) found that the APSIM-Wheat (with the model default frost parameters) could not account for extreme events such as severe frosts and might overestimate harvest yields under those conditions, based on an assessment of the model with data collected from the Wimmera region of Victoria, Australia. In 1998 the crops on one farm of this region were severely damaged by the stem frost and the model overestimated the harvest yield by more than 5 t/ha. Hochman et al. (2009) also reported an occurrence of both frost and heat damages in October 2004, late anthesis or early grain filling stages (the period when the crops are sensitive to extreme temperatures) in the Wimmera and Mallee regions of Victoria that caused the model to over-predict yield. For varieties with strong cold tolerance in the North China Plain, the minimum temperature threshold to cause leaf senescence was changed from -15°C to -20°C to eliminate the underestimation of LAI, biomass and yield (Chen et al., 2010b; Wang et al., 2009; Zhang et al., 2017, 2013). The modified temperature response of thermal time calculation and the temperature response of radiation use efficiency (RUE) led to further improve model simulations (Chen et al., 2010b, 2010c).

2.5.5.Other abiotic stresses

Some other factors APSIM-Wheat fails to simulate have been identified in model validation. The effect of soil cracking on soil evaporation is not taken into account in the reviewed model version, which leads the model to incorrectly simulate the water movement and further decreases yield prediction accuracy (Moeller et al., 2007; Mohanty et al., 2012; Paydar et al., 2005). Hochman et al. (2007) mentioned there was potential to improve yield prediction if a suitable function could be developed to describe the effects of subsoil constraints. Peake et al. (2014) tested the APSIM model performance in Queensland and New South Wales. During the wheat growth period, lodging was observed in most fields by the middle of grain filling and became worse after a large rain event. APSIM predicted lodging-impacted yield with a relatively large *RMSE*=3.26 t/ha, while non-lodged grain yield was predicted with *RMSE*=1.06 t/ha. When crops suffered hail damage in 1997 on one farm in the Wimmera region of Victoria, Australia, the model totally missed the hail storm damage and still predicted grain yield over 7 t/ha (Hochman et al., 2013). O'Leary et al. (2015) tested the APSIM-Wheat under two water regimes (irrigation and rain-fed), two nitrogen fertilisation regimes (0 and 53 – 138 kg N/ha), and two sowing dates for daytime ambient (365 µmol/mol) and elevated (550 µmol/mol) CO_2 environments at Horsham, Australia. The results indicated that the model showed a tendency to overestimate early biomass (DC31, stem elongation) (Zadoks et al., 1974), biomass at DC65 (anthesis), LAI at DC65 and grain yield under the normal CO_2 conditions; the resulting *RMSE* values were 1.592 t/ha, 1.542 t/ha, 0.70 m^2/m^2 and 1.294 t/ha, respectively. Under the elevated CO_2 condition, the model overcompensated the CO_2 effect and over predicted early biomass and harvest yields.

2.5.6.Biotic stress

Crops in most of the reviewed papers were well managed, with no significant insects, weeds, pests, or plant diseases observed. O'Leary et al. (2016) examined the performance of the APSIM-Wheat model under different stubble, tillage and nitrogen application management scenarios. Some large predictive errors were found when the model predicted yields for fields of Warwick, Australia, where the wheat was heavily infected with the root-lesion nematode. Biotic stress such as root disease load can have major impacts that are not represented in APSIM-Wheat yet. The simulated yield deviated more from the observed (*RMSE*=1.54 t/ha) when high nitrogen fertiliser was applied.

2.5.7.Implications of the influential factors in changing climate

Under future climate scenarios, both mean and variance of temperatures are projected to increase, along with precipitation variability. This will lead to increased heat waves, frost risk, and changing risk of drought and flood (Rigby and Porporato, 2008; Trenberth, 2011; Zeppel et al., 2014). The changing climate may also be favourable to certain wheat diseases, e.g. stripe rust (Luck et al., 2011) which have not been represented in the model yet. APSIM-Wheat, as a major cropping system tool used to study climate change impacts and seek solutions to address them (Deihimfard et al., 2018; Yang et al., 2014), needs improvement in the representation of heat stress, frost stress, water deficit, and the effects of pests, particularly when it is adopted to predict wheat production under the projected climate scenarios.

In addition to daily mean temperature, maximum temperature could be included as a variable to determine heat stress impact. The underestimation of heat stress impacts will lead to over-optimistic simulations of the future wheat production. Meanwhile, increasing mean temperatures accelerate crop growth and shorten the growing season, resulting in crops reaching the frost-sensitive anthesis stage more rapidly (Zheng et al., 2015). The absence of parameter values for functions to account for frosts can potentially lead to overestimation of harvest yields. Parameterising the frost damages of leaf senescence, seedling death, or death of formed grains will improve the model simulation capability. The variable precipitation intensity and probability may reduce users' confidence in simulation accuracy since the model showed uncertainty in predicting water-limited yield. Improved functioning and parametrisation to correctly estimate water deficit impacts on wheat growth is warranted. Apart from using the model to study future climate impacts on production, when users apply the model to a new study area or cultivar, accurate soil parameters and site-specifically calibrated cultivar parameters improve the model performance.

3. Meta-analysis of data from past studies

3.1. Building database for meta-analysis and performance metrics

All papers reviewed in Section 2 that had data that were extractable from tables, figures, text, or provided by the authors were included in the meta-database. In total, data from 30 studies were used to compose the meta-database for further analysis. These 30 studies are marked with asterisks in Table 3. Digitising the data from published scatter plots in the literature was performed with the WebPlotDigitizer tool (https://automeris.io/WebPlotDigitizer/). The database includes 1895 pairs of observed and simulated grain yields expressed in tons per ha. All these points were for validation simulations. The data originated from seven countries and included 51 wheat cultivars (see Table 3). These data were assembled and categorised according to different crop stresses and model initialisations. The conditions captured were:

- Crop stresses: water availability, nitrogen availability, heat stress, lodging, disease.
- Model initialisations: fully site-specific calibration, partially site-specific calibration, non-site-specific calibration.

APSIM Classic (model version please refer to Table 3) performance was evaluated for the whole data set and subsets corresponding to various conditions using the performance metrics in 2.2. To obtain R^2, a linear regression was fitted to the observed and simulated grain yield pairs. Residuals (simulated – observed yield) were also calculated and box plots drawn for different conditions. Comparisons between predicted yield residuals and observed yields were also plotted to visually investigate model capability and limitations. Statistics of coefficient of determination (R^2), *RMSE* (equation 3), *NRMSE* (equation 4), and *EF* (equation 5) were utilised to quantify the model performance.

3.2. Factors affecting APSIM yield prediction error

Overall, the model performed well. Figure 2 compares the predicted yield with the observed yield from the meta-database. APSIM-Wheat predicted grain yield with R^2 =0.68, *RMSE*=1.06 t/ha, *NRMSE*=28.89%, *EF*=0.63. This result is consistent with the findings from most papers reviewed in Section 2. To put these results in context of practical cropping decisions, Yield Prophet® users reported that discrepancies between the predicted and observed yields exceeding 0.5 t/ha reduced their confidence in using the model for decision support (Hochman et al., 2009a), indicating that factors contributing to the uncertainty and potential solutions should be explored. The deviation of observed vs. simulated yields scatters from the 1:1 line in Figure 2 (black dashed line) denotes model simulation deficiencies. The discrepancy between the regression line (grey dashed line) and the 1:1 line indicates existence of bias that varies from positive to negative values with yield. Potential causes of this bias include not fully site-specific calibration, water stress, nitrogen stress, heat stress, lodging, root-lesion nematode. The variation of yield prediction error and uncertainty under different environments, treatments, and model initialisations will be analysed in the following sections separately.

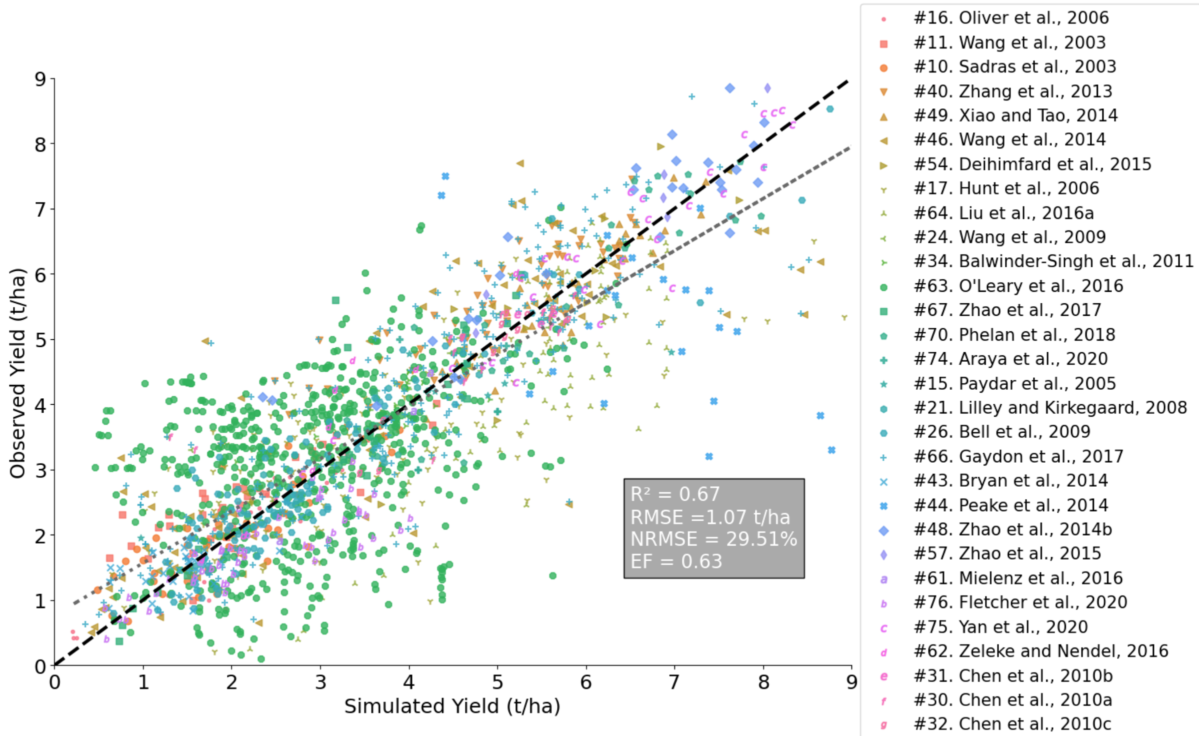

Figure 2. Comparison between observed and APSIM-Wheat simulated grain yields (black dashed line: 1:1 line; grey dashed line: regression line)

3.2.1.Site-specific calibration and specification of soil parameters

Figure 3 shows APSIM-Wheat validation results of the studies that used site-specific calibration. As described in Section 2.3, site-specific calibration is done by (1) manually tuning parameters to make the simulations correspond well with the observations or (2) specifying parameters with field measurements (usually soil texture, soil hydraulic and/or chemical parameters. The results indicated that the model, once site-specifically calibrated ((1), (2) individually or simultaneously), was able to estimate the harvest yield with an R^2 of 0.90, *RMSE*=0.64 t/ha, and a *NRMSE* of 14.08%. The model performance improved when model cultivar parameters were manually tuned and soil parameters were initialised with ground observations simultaneously (fully site-specific calibration), resulting in *RMSE* smaller than 0.5 t/ha, *NRMSE* of 10.15%, and an *EF* of 0.85, indicating that the model is performing well. When only the cultivar parameters were calibrated, the model maintained the *EF* of 0.8, with *RMSE* and *NRMSE* slightly increased to 0.7 t/ha and 12.93%, and an R^2 of 0.82. The yield prediction performance began to decline when only the soil parameters were specified with field measurements without adjusting other model parameters, both R^2 and *EF* decreased to 0.77, with an *NRMSE*=20.82%. The *RMSE* was only 0.51 t/ha since the yield range in this case were lower than in other cases.

Figure 4 shows results when soil parameters were specified using a soil database – APSoil or estimated soil hydraulic characteristics. Model default genotype parameters were utilised for specific cultivars. Compared to cases in Figure 3, these initialisation methods led to decrease the model accuracy and uncertainty, resulting in *RMSE* increasing from 0.64 to 1.25 t/ha and *NRMSE* increasing from 14.0% to 32.46%. When estimated soil hydraulic parameters were used, the *RMSE* of yield prediction was 1.37 t/ha and the *NRMSE* was 40.45%. Performance improved using the APSoil database to specify soil parameters resulting in model predictions with lower *RMSE* and *NRMSE* of 0.7 t/ha and 13.0%, compared with the model performance when using soil texture-derived soil parameters.

Results in Figure 3 and Figure 4 indicate that manually tuning cultivar parameters, and/or specifying the soil characteristics with ground observations can substantially improve the model performance. Convincing evidence is presented to demonstrate that APSIM-Wheat is able to simulate wheat grain yield within 0.5 t/ha when fully site-specific calibration implemented. When field measured data is not available, using a look-up-table approach that uses APSoil to specify the soil hydraulic properties can achieve yield prediction accuracy of *RMSE*=0.7 t/ha. Setting soil parameters with estimated data is still acceptable, but not ideal. The estimated soil parameters largely affect the yield prediction accuracy and uncertainty since they cannot appropriately describe the soil properties.

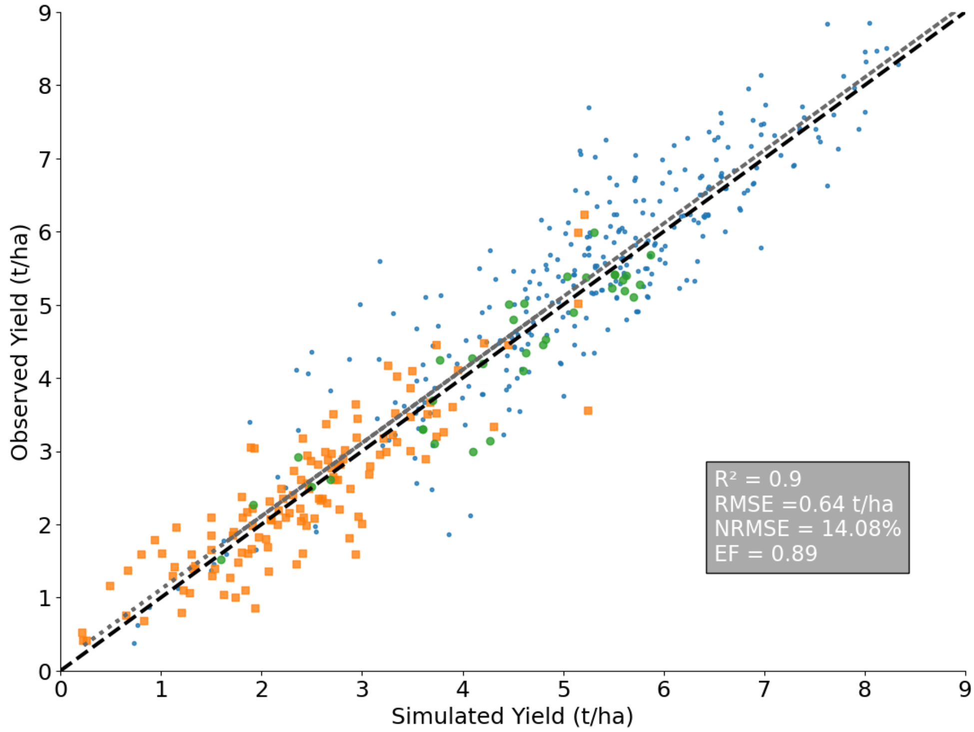

Figure 3. Comparison between observed and APSIM-Wheat simulated grain yields when cultivar parameters were manually tuned, and/or soil parameters were specified with ground observations (green circle: both cultivar and soil parameters were calibrated, R^2=0.87, RMSE=0.44 t/ha, NRMSE=10.15%, EF=0.85; blue dot: only cultivar parameters were tuned, R^2=0.82, RMSE=0.7 t/ha, NRMSE=12.93%, EF=0.8; orange square: only soil parameters were specified using field measurements, R^2=0.77, RMSE=0.51 t/ha, NRMSE=20.82%, EF=0.77)

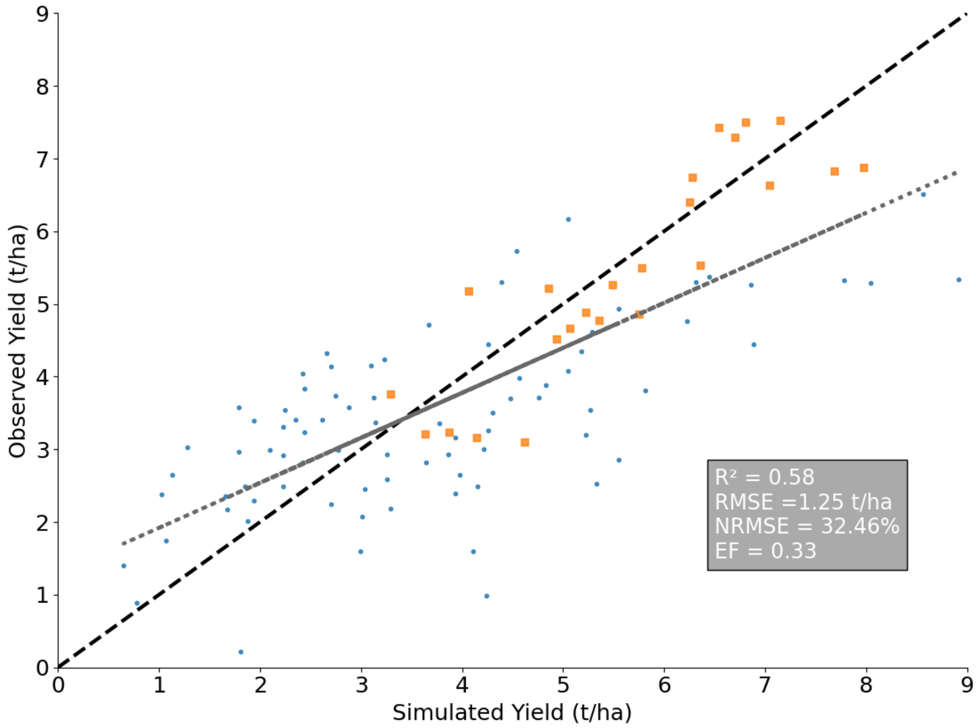

Figure 4. Comparison between observed and APSIM-Wheat simulated grain yields when soil parameters specified using APSoil or estimated data (blue dot: estimated soil characteristics,

R^2=0.45, RMSE=1.37 t/ha, NRMSE=40.45%, EF=-0.27; orange square: soil parameters were specified using APSoil, R^2=0.78, RMSE=0.7 t/ha, NRMSE=13.0%, EF=0.76)

3.2.2. Water availability

To assess the impacts of water availability on APSIM yield prediction, only site-specifically calibrated datasets from irrigated or water limited fields have been selected. Figure 5 shows box plots of prediction residuals. The cases are presented in the order of water stress, from highest (1) to lowest (3). Case 1 shows datasets for crops under water limited conditions. Datasets from two papers were included (Fletcher et al., 2020; Peake et al., 2014). Wheat from Case 2 was irrigated at critical growth stages with different amounts of water (Balwinder-Singh et al., 2011; Chen et al., 2010b; Deihimfard et al., 2015; Wang et al., 2013; Xiao and Tao, 2014; Yan et al., 2020; Zhang et al., 2013; Zhao et al., 2014b). Wheat from Case 3 was also irrigated, but not at specific growth stages. The irrigation amount and scheduling were adapted to the actual water demand (Gaydon et al., 2017; Mielenz et al., 2016; Yan et al., 2020).

Wheat in Case 1 suffered from water stress. Peake et al. (2014) observed mild water stress during the pre- and post-anthesis, while the model was also used to predict water-limited yield (Fletcher et al., 2020). 50% of the predicted yield residuals were within the range of 0.2–1 t/ha, 99.3% of them were within the range -0.4–2.4 t/ha, while the median was approximately 0.33 t/ha. From the datasets we analysed, yield overestimation was more obvious than underestimation under water stressed conditions. In Case 2, the fields were mainly from Punjab, India, North China Plain (NCP) and North-eastern Iran. They were irrigated at critical growth stages, e.g., sowing, jointing, flowering and grain filling, with total irrigation amounts between 75 and 450 mm. The accuracy of modelled yields was acceptable with *RMSE* around 0.65 t/ha. The median of residuals of modelled yields did not exceed -0.5 t/ha. Approximately 50% of the predicted yield residuals were within the range of -0.65–0.15 t/ha, and 99.3% of them were within the range of -1.7–1.25 t/ha. Underestimation was more obvious than overestimation. Case 3 shows crops irrigated according to their water demand. Irrigation scheduling and amount were adjusted according to rainfall amount, soil water content, and crop requirement. Crops in this case barely experienced water limitation and the model performance was more accurate and stable. The residual medians were less than 0.2 t/ha, and 99.3% of the residuals were within ±0.7 t/ha.

Case 1 demonstrated that APSIM-Wheat tend to overestimate yield with more significant uncertainty under water-limited conditions. It seems that the constraint on wheat growth by limited water is not well accounted for by APSIM-Wheat, leading to overly optimistic grain yield prediction. The mechanism that APSIM-Wheat uses to handle water stress was described in Section 2.1. The model only accounts for water deficit impacts on biomass production and leaf expansion. It does include a function intended to account for water stress on phenology, but the default parameterisation results in no effect on phenology. Consequently, proper parametrisation to correctly estimate drought impact on phenology under water-limited condition is needed. For example, Chauchan et al. (2019) accounted for soil water effect to modulate APSIM Classic (version 7.10) predicted flowering time. But the proposed method can only reduce the daily thermal time and delay the flowering time when the soil water is sufficient (fractional available soil water>0.65). A proper scheme to directly simulate the impact of soil water stress on flowering time is yet to be developed. Greater water limitations result in higher canopy temperatures, which reduce the duration of biomass accumulation. The increased canopy temperatures under water deficit conditions should be considered to improve the performance of yield prediction (Asseng et al., 2004). Case 2 showed that with critical-stage irrigation the model can predict yield with acceptable accuracy (*RMSE*=0.65 t/ha), while the uncertainty is still obvious. These datasets demonstrated that once extra water was supplied (in addition to rainfall), APSIM-Wheat could capture the additional water resource. Case 3 showed that supplying irrigation water according to crop demand to avoid water limitation was associated with better modelling performance.

Users operating APSIM in a water-limited situation should be aware of uncertainty and possible yield overestimation. Most researchers validate the model using real-world datasets to create confidence in its performance before using it in combination with climate projections for predicting food production under climate change scenarios. The frequency and intensity of droughts are projected to increase (Zhou et al., 2019) and the water availability for rain-fed agriculture is decreasing, and the crop model will probably underestimate yields under those conditions. Larger prediction uncertainty should be considered when utilising cropping system as a tool to assess future food production and security.

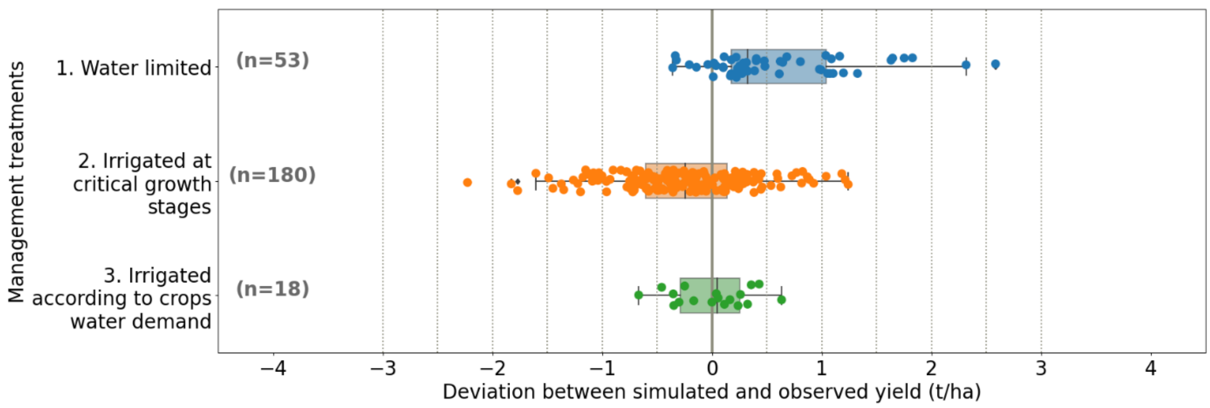

Figure 5. Boxplot of APSIM predicted yield residuals under different irrigation practices

3.2.3.Nitrogen availability

We selected site-specific calibrated datasets to assess the impacts of nitrogen availability on APSIM yield prediction, in the absence of other stresses. Figure 6 shows box plots of prediction residuals of six cases, which were ordered from the largest to the smallest nitrogen stress. Case 1 shows datasets when crops experience nitrogen limitation. Datasets from two papers were included (Peake et al., 2014; Wang et al., 2014). Case 2 was also composed of datasets from two papers (Sadras et al., 2003; Zhao et al., 2014b). The authors did not specify the nitrogen rate in these datasets but declared no nitrogen stress was observed. Wheat from Cases 3–6 were fertilised with different rates of nitrogen. The application amount increased from 64 kg N/ha to 195 kg N/ha. Data for Case 3 were collected from three papers (Araya et al., 2020; Paydar et al., 2005; Phelan et al., 2018), while Cases 4–6 used datasets from Xiao and Tao (2014) and Yan et al. (2020).

Case 1 reported nitrogen stress symptoms (leaf yellowing) at DC31 (early stem elongation) (Peake et al., 2014), while the model was also used to predict yield when no fertiliser was applied in fields (Wang et al., 2014). Both overestimation and underestimation were observed with the median of residuals approximately -0.3 t/ha. In some cases, the underestimation was even more than 3 t/ha. Case 2 collected datasets with not specified fertilisation amounts, but no nitrogen stress was observed in the fields. The model predicted yield with acceptable accuracy and uncertainty. The median of residuals was close to zero. 50% of the predicted yield residuals were within the range of -0.5–0.2 t/ha, and 99.3% of them were within the range of -1.25–1 t/ha. Case 3 contained datasets with nitrogen rates of 64, 75, 100, and 128 kg N/ha. The distribution of the prediction residuals was similar to those in Case 2. With the increasing nitrogen application rate, the predicted yield residuals were less scattered, ranging within ±1.0 t/ha to ±0.5 t/ha while the medians tended towards 0 t/ha. The model was well-performed to catch the fertilisation differences.

Under nitrogen limited conditions (Case 1), APSIM-Wheat showed the largest uncertainties and more severe yield underestimation and the model tended to underestimate yield when crops suffered from nitrogen limitation. The reason as indicated by Peake et al. (2014) is that APSIM-Wheat overrated the nitrogen stress duration by two weeks longer compared to observed nitrogen stress in the paddocks. Cases 2–6 showed that once extra nitrogen was supplied, the model captured the increasing trend and tended to predict yield with better accuracy and lower uncertainty. The residuals were contained within ±1.0 t/ha when sufficient nitrogen was applied.

An additional parametrisation of the nitrogen impacts on phenology would be able to better address potential simulation problems. Zhao et al. (2014a) assessed the nitrogen concentration parameters used in the model, the results indicated that the higher leaf maximum and critical nitrogen concentrations led the model to overrate the nitrogen stress impacts on biomass accumulation and underrate the impacts on leaf expansion. They suggested to adjust and verify these parameters to increase the prediction accuracy of the model for grain yield.

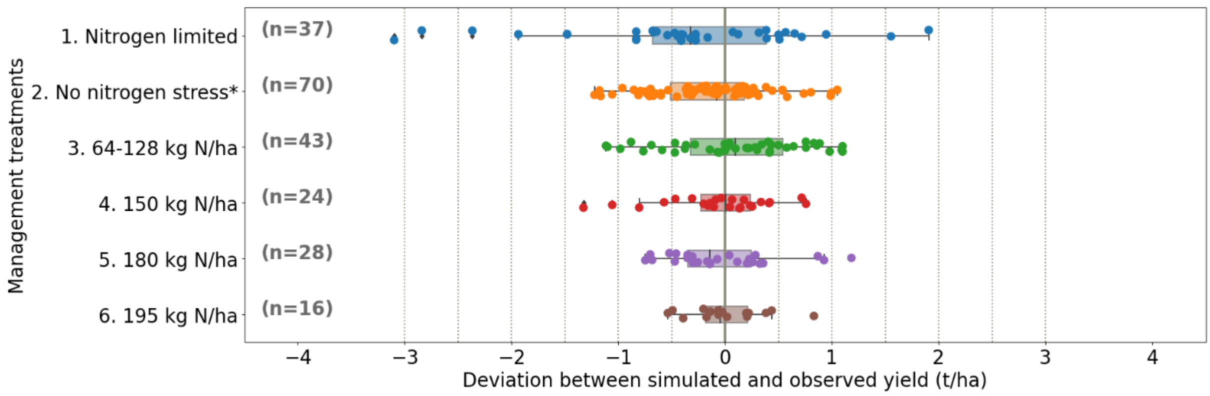

Figure 6. Boxplot of APSIM predicted yield residuals under different nitrogen application rates (* fertiliser amount was not specified)

3.2.4. Other stresses

Figure 7 illustrates the model predicted yield residuals against the observed yield for datasets under irrigated and fertilised conditions. We intentionally included datasets for wheat without stress and under some abiotic stresses such as heat and lodging, to compare the model performance under stressed and stress-free situations (Deihimfard et al., 2015; Liu et al., 2016b; Mielenz et al., 2016; Peake et al., 2014; Xiao and Tao, 2014; Yan et al., 2020; Zeleke and Nendel, 2016; Zhao et al., 2015, 2014b). The model showed a good performance for all stress-free cases, with *RMSE*=0.66 t/ha and *NRMSE*=12.49%. However, when the stressed cases are included, *RMSE* increased to around 1.03 t/ha and *NRMSE* to 20.26%, respectively. The mean residual is 0.3 t/ha and standard deviation is 0.99 t/ha. Most of the residuals are between the range of $\pm$ 1.96 times of the standard deviation around the mean. The outliers are from the cases where the crops were under heat stress and impacted by lodging.

Heat stress. Figure 7 indicates that the model cannot capture well the effects from a short-term exposure of wheat to extreme high temperatures, especially during anthesis and grain filling stages (orange squares and green circles). The datasets were from a environment-controlled phytotron experiment (Liu et al., 2016a). In these cases, yields are over-predicted, especially if the heat stress occurred during anthesis with *RMSE*=1.5 t/ha and *NRMSE*=35.47%, while heat stress during grain filling resulted in *RMSE*=1.14 t/ha and *NRMSE*=22.75%. Barlow et al. (2015) emphasised the need to define response functions for calculating extreme temperatures impacts, with a priority on the response during anthesis and grain filling stages.

Lodging. The brown triangular points from Figure 7 represent the model predicted residuals against the observed yield when crops were impacted by lodging (the data is from Peake et al. (2014)). Yield is severely over-estimated with *RMSE*=3.26 t/ha, *NRMSE*=76.77%. The reviewed APSIM-Wheat version does not consider effects of crop lodging, while lodging can be caused by many factors, e.g., excessive nitrogen fertilisation and irrigation, heavy rain, wind, or hailstorm. The development of functions in

APSIM-Wheat that accounts for the effects of lodging would be desirable although it would require collection of extensive databases of crops affected by lodging.

Our results suggest that the projected increasing frequency, intensity and duration of global heat waves (Perkins et al., 2012), extreme weather events (Meehl et al., 2000), and floods (Kundzewicz et al., 2014) can lead to greater uncertainties in the simulation of future climate scenarios with APSIM-Wheat. However, as mentioned in Section 3.2.2, model users tend to trust the model performance even when the model is applied to project crop yield under future scenarios, which could lead to overly optimistic food production by underestimating the negative effects of heat and lodging.

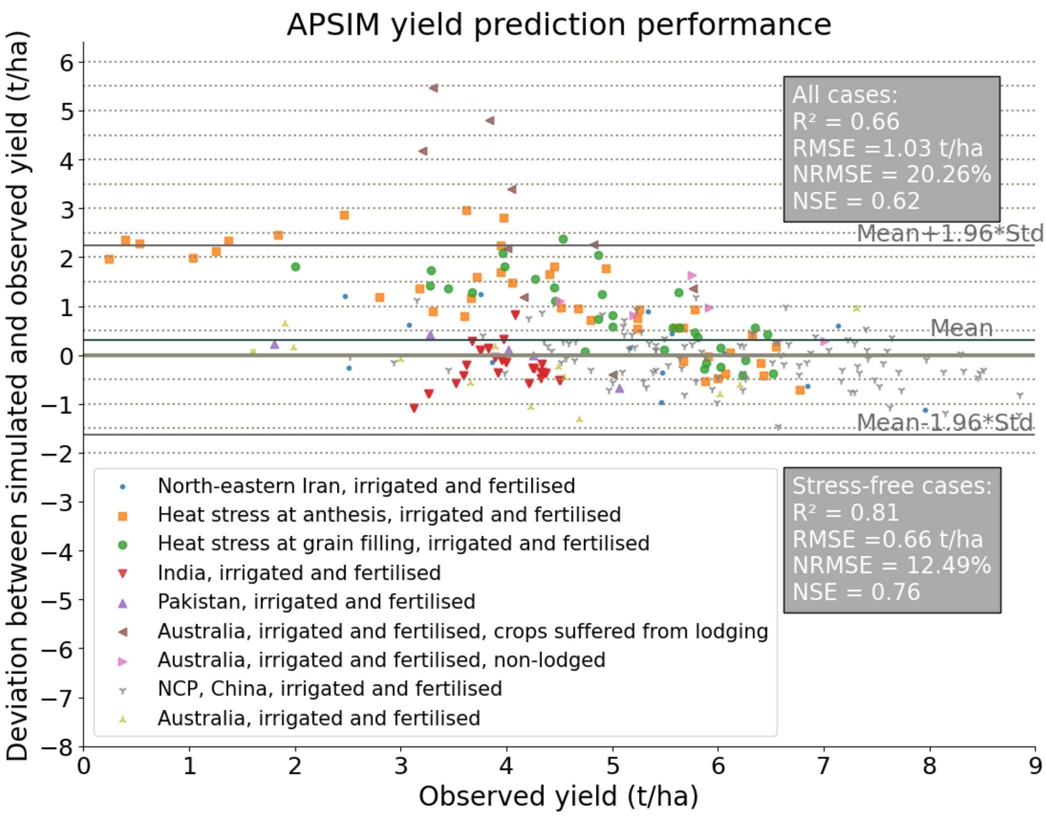

Figure 7. Comparison between predicted yield residuals and observed yield under irrigated and fertilised condition

4. Summary and Conclusion

In this work, we have reviewed 76 articles and conducted a meta-analysis of 30 applications of the APSIM model (APSIM Classic, version 1.X – version 7.9) to obtain detailed information on the process-based model's performance in predicting wheat yields. Our study shows that the model provides reasonably accurate wheat grain yields across a wide range of varieties, environments and management practices around the world with an overall uncertainty of about 1 t/ha. However, we found a large variation in uncertainties within the modelling studies considered, especially between studies with site-specific calibration and non-site-specific calibration.

Furthermore, we found that factors such as heat and frost stress, water and nitrogen availability, soil parameterisation, calibration of genotype parameters, soil cracking, lodging, increased atmospheric CO_2 concentration and plant diseases are important factors affecting model performance. Heat and frost stresses, in particular, caused large discrepancies in the prediction of grain yield. One reason for this is that the reviewed model versions use only daily mean temperature as a heat factor to calculate the effects on biomass accumulation, although wheat is particularly sensitive to shorter-term heat stress during the anthesis and grain filling phases. Therefore, APSIM tended to overestimate crop yield that experienced heat wave conditions. Frost stress functions are already implemented in the reviewed model but without default parametrisation which negates their effect (impact factor = 0) so APSIM overestimates yield in crops subject to frost damage. The applications of APSIM to situations with water stress and nitrogen limitation led to greater uncertainties (overestimation for water stress and underestimation for nitrogen stress). Like the frost stress function, the effects of water and nitrogen stress on phenology are not yet parameterised.

A fully or partial site-specific calibration resulted in crop yields being predicted with higher accuracy (on average, *RMSE* and *NRMSE* were 0.64 t/ha and 14.08%, respectively). A fully site-specific calibration, including the determination of soil hydraulic parameters, initial soil conditions from field measurements and adjustment of other parameters (such as crop parameters), resulted in the lowest uncertainty in crop prediction (*RMSE*=0.44 t/ha, *NRMSE*=10.15%). If soil parameters are not available, using a soil database such as APSoil to specify soil hydraulic properties is a good alternative, leading to yield predictions on average with *RMSE*=0.7 t/ha and *NRMSE*=13.0%. Soil texture-derived soil parameterisation is also acceptable but with comparatively lower accuracy and uncertainty with an *RMSE*=1.37 t/ha and an *NRMSE*=40.45%.

The reviewed APSIM-Wheat version is not equipped with functions that account for other abiotic and biotic influences like soil cracking, lodging, or crop disease. Improving the model functionally to consider all these factors could lead to better crop predictions; however, a major challenge is that there is often a significant stochastic component to these influences. An alternative would be to pursue methods such as assimilating external observations into the model to continuously adjust certain model state variables and properties to improve model performance. Remote sensing data can provide timely information on the crop or environment status and could be used to update the model simulation regularly during the simulation. Another option is to use multi-model ensembles to account for model uncertainty in describing the impact of climate change on agricultural productivity (Asseng et al., 2015, 2013; Iizumi et al., 2018; Maiorano et al., 2017; Martre et al., 2015b; Wang et al., 2017).

This work did not assess the model's ability of simulating other crop states such as biomass, leaf area, water use, or fertility dynamics. The simulation quality of these dynamics is still largely unknown and worth further investigation.

The meta-database in this paper was composed of datasets from separate papers. In our meta-analysis, datasets from existing papers were compiled to analyse the impact of certain factors, while other factors could not be held constant, which may have led to some bias. The model validations considered in our study were all point-based, the plant models are usually used at the plot, field scales or even larger. However, the effects of spatial heterogeneity were not considered in our study. Finally, we did not consider the uncertainties embedded in the forcing inputs. According to Tao et al. (2018), when coupling climate models with crop models, the uncertainty from downscaled climate projections could be larger than those from crop models.

Crop models like APSIM are not just predictive tools, but also exploratory tools in conjunction with future scenarios. The vision for agricultural systems models is to accelerate progress of finding ways to address the global food security challenges. This paper aims to provide the perspectives on the model outputs credibility and uncertainty under various conditions covering a wide spectrum of management practices, environments and wheat varieties. We expect that our analysis of APSIM-Wheat model performance will assist users to have appropriate interpretations and avoid misuse of the model.

Declaration of Competing Interest

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

Acknowledgements

The authors would like to thank Garry J. O'Leary (Agriculture Victoria, Department of Jobs, Precincts and Regions) and Keith G. Pembleton (University of Southern Queensland) for providing valuable and insightful comments that improved the quality of this paper. We gratefully acknowledge Donald S. Gaydon (CSIRO Agriculture and Food, Australia), Peter J. Innes, Hamish E. Brown (New Zealand Institute for Plant and Food Research) and Neil I. Huth (CSIRO Agriculture and Food, Australia) for providing the datasets used in the evaluation of the model; Peter Carberry, Zvi Hochman, and Victor Sadras for friendly replies and suggestions. Harrie-Jan Hendricks Franssen acknowledges support from the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy, EXC-2070–390732324–PhenoRob. Acknowledgement is also made to the APSIM Initiative which takes responsibility for quality assurance and a structured innovation programme for APSIM's modelling software, which is provided free for research and development use (see www.apsim.info for details).

Reference

Acuña, T.B., Lisson, S., Johnson, P., Dean, G., 2015. Yield and water-use efficiency of wheat in a high-rainfall environment. Crop Pasture Sci. 66, 419–429.

Ahmed, M., Akram, M.N., Asim, M., Aslam, M., ul Hassan, F., Higgins, S., Stöckle, C.O., Hoogenboom, G., 2016. Calibration and validation of APSIM-Wheat and CERES-Wheat for spring wheat under rainfed conditions: Models evaluation and application. Comput. Electron. Agric. 123, 384–401. https://doi.org/10.1016/j.compag.2016.03.015

Akinseye, F.M., Ajeigbe, H.A., Traore, P.C.S., Agele, S.O., Zemadim, B., Whitbread, A., 2020. Improving sorghum productivity under changing climatic conditions: A modelling approach. F. Crop. Res. 246, 107685.

Anwar, M.R., O'Leary, G.J., Rab, M.A., Fisher, P.D., Armstrong, R.D., 2009. Advances in precision agriculture in south-eastern Australia. V. Effect of seasonal conditions on wheat and barley yield response to applied nitrogen across management zones. Crop Pasture Sci. 60, 901–911.

Anwar, M.R., Wang, B., Li Liu, D., Waters, C., 2020. Late planting has great potential to mitigate the effects of future climate change on Australian rain-fed cotton. Sci. Total Environ. 714, 136806.

Araya, A., Kisekka, I., Girma, A., Hadgu, K.M., Tegebu, F.N., Kassa, A.H., Ferreira-Filho, H.R., Beltrao, N.E., Afewerk, A., Abadi, B., others, 2017. The challenges and opportunities for wheat production under future climate in Northern Ethiopia. J. Agric. Sci. 155, 379–393.

Araya, A., Prasad, P.V. V, Gowda, P.H., Djanaguiraman, M., Kassa, A.H., 2020. Potential impacts of climate change factors and agronomic adaptation strategies on wheat yields in central highlands of Ethiopia. Clim. Change 1–19.

Archontoulis, S. V, Miguez, F.E., Moore, K.J., 2014. Evaluating APSIM maize, soil water, soil nitrogen, manure, and soil temperature modules in the Midwestern United States. Agron. J. 106, 1025–1040.

Asseng, S., Anderson, G.C., Dunin, F.X., Fillery, I.R.P., Dolling, P.J., Keating, B.A., 1998a. Use of the APSIM wheat model to predict yield, drainage, and NO_3-leaching for a deep sand. Aust. J. Agric. Res. 49, 363–378.

Asseng, S., Bar-Tal, A., Bowden, J.W., Keating, B.A., Van Herwaarden, A., Palta, J.A., Huth, N.I., Probert, M.E., 2002. Simulation of grain protein content with APSIM-Nwheat. Eur. J. Agron. 16, 25–42.

Asseng, S., Ewert, F., Martre, P., Rötter, R.P., Lobell, D.B., Cammarano, D., Kimball, B.A., Ottman, M.J., Wall, G.W., White, J.W., Reynolds, M.P., Alderman, P.D., Prasad, P.V.V., Aggarwal, P.K., Anothai, J., Basso, B., Biernath, C., Challinor, A.J., De Sanctis, G., Doltra, J., Fereres, E., Garcia-Vila, M., Gayler, S., Hoogenboom, G., Hunt, L.A., Izaurralde, R.C., Jabloun, M., Jones, C.D., Kersebaum, K.C., Koehler, A.K., Müller, C., Naresh Kumar, S., Nendel, C., O'leary, G., Olesen, J.E., Palosuo, T., Priesack, E., Eyshi Rezaei, E., Ruane, A.C., Semenov, M.A., Shcherbak, I., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P.J., Waha, K., Wang, E., Wallach, D., Wolf, J., Zhao, Z., Zhu, Y., 2015. Rising temperatures reduce global wheat production. Nat. Clim. Chang. 5, 143–147. https://doi.org/10.1038/nclimate2470

Asseng, S., Ewert, F., Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P.J., Rötter, R.P., Cammarano, D., Brisson, N., Basso, B., Martre, P., Aggarwal, P.K., Angulo, C., Bertuzzi, P., Biernath, C., Challinor, A.J., Doltra, J., Gayler, S., Goldberg, R., Grant, R., Heng, L., Hooker, J., Hunt, L.A., Ingwersen, J., Izaurralde, R.C., Kersebaum, K.C., Müller, C., Naresh Kumar, S., Nendel, C., O'Leary, G., Olesen, J.E., Osborne, T.M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M.A., Shcherbak, I., Steduto, P., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., Wallach, D., White, J.W., Williams, J.R., Wolf, J., 2013. Uncertainty in simulating wheat yields under climate change. Nat. Clim. Chang. 3, 827–832. https://doi.org/10.1038/nclimate1916

Asseng, S., Foster, I.A.N., Turner, N.C., 2011. The impact of temperature variability on wheat yields. Glob. Chang. Biol. 17, 997–1012.

Asseng, S., Jamieson, P.D., Kimball, B., Pinter, P., Sayre, K., Bowden, J.W., Howden, S.M., 2004. Simulated wheat growth affected by rising temperature, increased water deficit and elevated atmospheric CO_2. F. Crop. Res. 85, 85–102. https://doi.org/10.1016/S0378-4290(03)00154-0

Asseng, S., Keating, B.A., Fillery, I.R.P., Gregory, P.J., Bowden, J.W., Turner, N.C., Palta, J.A., Abrecht, D.G., 1998b. Performance of the APSIM-wheat model in Western Australia. F. Crop. Res. 57, 163–179.

Asseng, S., Turner, N.C., Keating, B.A., 2001. Analysis of water-and nitrogen-use efficiency of wheat in a Mediterranean climate. Plant Soil 233, 127–143.

Asseng, S., Van Keulen, H., Stol, W., 2000. Performance and application of the APSIM Nwheat model in the Netherlands. Eur. J. Agron. 12, 37–54.

Bahri, H., Annabi, M., M'Hamed, H.C., Frija, A., 2019. Assessing the long-term impact of conservation agriculture on wheat-based systems in Tunisia using APSIM simulations under a climate change context. Sci. Total Environ. 692, 1223–1233.

Bai, H., Wang, J., Fang, Q., Huang, B., 2020. Does a trade-off between yield and efficiency reduce water and nitrogen inputs of winter wheat in the North China Plain? Agric. Water Manag. 233, 106095.

Balwinder-Singh, Gaydon, D.S., Humphreys, E., Eberbach, P.L., 2011. The effects of mulch and irrigation management on wheat in Punjab, India—Evaluation of the APSIM model. F. Crop. Res. 124, 1–13.

Barlow, K.M., Christy, B.P., O'leary, G.J., Riffkin, P.A., Nuttall, J.G., 2015. Simulating the impact of extreme heat and frost events on wheat crop production: A review. F. Crop. Res. 171, 109–119.

Bell, L.W., Hargreaves, J.N.G., Lawes, R.A., Robertson, M.J., 2009. Sacrificial grazing of wheat crops: identifying tactics and opportunities in Western Australia's grainbelt using simulation approaches. Anim. Prod. Sci. 49, 797–806.

Bosi, C., Sentelhas, P.C., Huth, N.I., Pezzopane, J.R.M., Andreucci, M.P., Santos, P.M., 2020. APSIM-Tropical Pasture: A model for simulating perennial tropical grass growth and its parameterisation for palisade grass (Brachiaria brizantha). Agric. Syst. 184, 102917.

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., others, 2003. An overview of the crop model STICS. Eur. J. Agron. 18, 309–332.

Brisson, N., Mary, B., Ripoche, D., Jeuffroy, M.H., Ruget, F., Nicoullaud, B., Gate, P., Devienne-Barret, F., Antonioletti, R., Durr, C., others, 1998. STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn.

Brisson, N., Ruget, F., Gate, P., Lorgeou, J., Nicoullaud, B., Tayot, X., Plenet, D., Jeuffroy, M.-H., Bouthier, A., Ripoche, D., others, 2002. STICS: a generic model for simulating crops and their water and nitrogen balances. II. Model validation for wheat and maize. Agronomie 22, 69–92.

Brown, H., Huth, N., Holzworth, D., 2018. Crop model improvement in APSIM: Using wheat as a case study. Eur. J. Agron. 0–1. https://doi.org/10.1016/j.eja.2018.02.002

Brown, H.E., Huth, N.I., Holzworth, D.P., Teixeira, E.I., Zyskowski, R.F., Hargreaves, J.N.G., Moot, D.J., 2014. Plant modelling framework: software for building and running crop models on the APSIM platform. Environ. Model. Softw. 62, 385–398.

Bryan, B.A., King, D., Zhao, G., 2014. Influence of management and environment on Australian wheat: information for sustainable intensification and closing yield gaps. Environ. Res. Lett. 9, 44005.

Carberry, P.S., Hochman, Z., Hunt, J.R., Dalgliesh, N.P., McCown, R.L., Whish, J.P.M., Robertson, M.J., Foale, M.A., Poulton, P.L., Van Rees, H., 2009. Re-inventing model-based decision support with Australian dryland farmers. 3. Relevance of APSIM to commercial crops. Crop Pasture Sci. 60, 1044–1056.

Carberry, P.S., Liang, W., Twomlow, S., Holzworth, D.P., Dimes, J.P., McClelland, T., Huth, N.I., Chen, F., Hochman, Z., Keating, B.A., 2013. Scope for improved eco-efficiency varies among diverse cropping systems. Proc. Natl. Acad. Sci. 110, 8381–8386.

Casadebaig, P., Zheng, B., Chapman, S., Huth, N., Faivre, R., Chenu, K., 2016. Assessment of the potential impacts of wheat plant traits across environments by combining crop modeling and global sensitivity analysis. PLoS One 11, e0146385.

Chauhan, Y.S., Ryan, M., Chandra, S., Sadras, V.O., 2019. Accounting for soil moisture improves prediction of flowering time in chickpea and wheat. Sci. Rep. 9, 1–11.

Chen, C., Wang, E., Yu, Q., 2010a. Modelling the effects of climate variability and water management on crop water productivity and water balance in the North China Plain. Agric. Water Manag. 97, 1175–1184. https://doi.org/10.1016/j.agwat.2008.11.012

Chen, C., Wang, E., Yu, Q., 2010b. Modeling wheat and maize productivity as affected by climate variation and irrigation supply in North China Plain. Agron. J. 102, 1037–1049.

Chen, C., Wang, E., Yu, Q., Zhang, Y., 2010c. Quantifying the effects of climate trends in the past 43 years (1961--2003) on crop growth and water demand in the North China Plain. Clim. Change 100, 559–578.

Connolly, R.D., Bell, M., Huth, N., Freebairn, D.M., Thomas, G., 2002. Simulating infiltration and the water balance in cropping systems with APSIM-SWIM. Soil Res. 40, 221–242.

Dalgliesh, N., Cocks, B., Horan, H., others, 2012. APSoil-providing soils information to consultants, farmers and researchers, in: 16th Australian Agronomy Conference, Armidale, NSW.

Dalgliesh, N.P., Foale, M.A., McCown, R.L., 2009. Re-inventing model-based decision support with Australian dryland farmers. 2. Pragmatic provision of soil information for paddock-specific simulation and farmer decision making. Crop Pasture Sci. 60, 1031–1043.

Deihimfard, R., Eyni-Nargeseh, H., Mokhtassi-Bidgoli, A., 2018. Effect of future climate change on wheat yield and water use efficiency under semi-arid conditions as predicted by APSIM-wheat model. Int. J. Plant Prod. 12, 115–125.

Deihimfard, R., Mahallati, M.N., Koocheki, A., 2015. Yield gap analysis in major wheat growing areas of Khorasan province, Iran, through crop modelling. F. Crop. Res. 184, 28–38.

FAO, 2020. World Food and Agriculture - Statistical Yearbook 2020. Rome.

Fisher, J.S., Asseng, S., Bowden, J.W., Robertson, M.J., others, 2001. Trial-years without tears: enhancing recommendations of flowering and yield in wheat, in: Proceedings of the 10th Australian Agronomy Conference, Hobart,(Australian Society of Agronomy). Http://Www. Regional. Org. Au/Au/Asa/2001/6/d/Fisher. Htm.

Fletcher, A.L., Chen, C., Ota, N., Lawes, R.A., Oliver, Y.M., 2020. Has historic climate change affected the spatial distribution of water-limited wheat yield across Western Australia? Clim. Change 1–18.

Frederiks, T.M., Christopher, J.T., Harvey, G.L., Sutherland, M.W., Borrell, A.K., 2012. Current and emerging screening methods to identify post-head-emergence frost adaptation in wheat and barley. J. Exp. Bot. 63, 5405–5416.

Fuller, M.P., Fuller, A.M., Kaniouras, S., Christophers, J., Fredericks, T., 2007. The freezing characteristics of wheat at ear emergence. Eur. J. Agron. 26, 435–441.

Gaydon, D.S., Lisson, S.N., Xevi, E., others, 2006. Application of APSIM 'multi-paddock'to estimate whole-of-farm water-use efficiency, system water balance and crop production for a rice-based operation in the Coleambally Irrigation District, NSW, in: Proceedings of the 13th Australian Society of Agronomy Conference. pp. 10–14.

Gaydon, D.S., Wang, E., Poulton, P.L., Ahmad, B., Ahmed, F., Akhter, S., Ali, I., Amarasingha, R., Chaki, A.K., Chen, C., others, 2017. Evaluation of the APSIM model in cropping systems of Asia. F. Crop. Res. 204, 52–75.

Hammer, G.L., van Oosterom, E., McLean, G., Chapman, S.C., Broad, I., Harland, P., Muchow, R.C., 2010. Adapting APSIM to model the physiology and genetics of complex adaptive traits in field crops. J. Exp. Bot. 61, 2185–2202.

Hansen, J.W., Mishra, A., Rao, K.P.C., Indeje, M., Ngugi, R.K., 2009. Potential value of GCM-based seasonal rainfall forecasts for maize management in semi-arid Kenya. Agric. Syst. 101, 80–90.

He, L., Cleverly, J., Chen, C., Yang, X., Li, J., Liu, W., Yu, Q., 2014. Diverse responses of winter wheat yield and water use to climate change and variability on the semiarid Loess Plateau in China. Agron. J. 106, 1169–1178.

Hochman, Z., Dang, Y.P., Schwenke, G.D., Dalgliesh, N.P., Routley, R., McDonald, M., Daniells, I.G., Manning, W., Poulton, P.L., 2007. Simulating the effects of saline and sodic subsoils on wheat crops growing on Vertosols. Aust. J. Agric. Res. 58, 802–810.

Hochman, Z., Gobbett, D., Holzworth, D., McClelland, T., van Rees, H., Marinoni, O., Garcia, J.N., Horan, H., 2013. Reprint of "Quantifying yield gaps in rainfed cropping systems: A case study of wheat in Australia." F. Crop. Res. 143, 65–75.

Hochman, Z., Holzworth, D., Hunt, J.R., 2009a. Potential to improve on-farm wheat yield and WUE in Australia. Crop Pasture Sci. 60, 708–716.

Hochman, Z., Van Rees, H., Carberry, P.S., Hunt, J.R., McCown, R.L., Gartmann, A., Holzworth, D., Van Rees, S., Dalgliesh, N.P., Long, W., others, 2009b. Re-inventing model-based decision support with Australian dryland farmers. 4. Yield Prophet®helps farmers monitor and manage crops in a variable climate. Crop Pasture Sci. 60, 1057–1070.

Holzworth, D., Huth, N.I., Fainges, J., Brown, H., Zurcher, E., Cichota, R., Verrall, S., Herrmann, N.I., Zheng, B., Snow, V., 2018. APSIM Next Generation: Overcoming challenges in modernising a farming systems model. Environ. Model. Softw. 103, 43–51.

Holzworth, D.P., Huth, N.I., deVoil, P.G., 2011. Simple software processes and tests improve the reliability and usefulness of a model. Environ. Model. Softw. 26, 510–516.

Holzworth, D.P., Huth, N.I., DeVoil, P.G., Zurcher, E.J., Herrmann, N.I., McLean, G., Chenu, K., van Oosterom, E.J., Snow, V., Murphy, C., others, 2014. APSIM--evolution towards a new generation of agricultural systems simulation. Environ. Model. Softw. 62, 327–350.

Horie, T., Yajima, M., Nakagawa, H., 1992. Yield forecasting. Agric. Syst. 40, 211–236.

Hsiao, T.C., Heng, L., Steduto, P., Rojas-Lara, B., Raes, D., Fereres, E., 2009. AquaCrop—The FAO crop model to simulate yield response to water: III. Parameterization and testing for maize. Agron. J. 101, 448–459.

Hunt, J., van Rees, H., Hochman, Z., Carberry, P., Holzworth, D., Dalgliesh, N., Brennan, L., Poulton, P., van Rees, S., Huth, N.I., others, 2006. Yield Prophet®: An online crop simulation service, in: Proceedings of the 13th Australian Agronomy Conference. pp. 10–14.

Hussain, J., Khaliq, T., Ahmad, A., Akhtar, J., 2018. Performance of four crop model for simulations of wheat phenology, leaf growth, biomass and yield across planting dates. PLoS One 13, e0197546.

Iizumi, T., Shin, Y., Kim, W., Kim, M., Choi, J., 2018. Global crop yield forecasting using seasonal climate information from a multi-model ensemble. Clim. Serv. 11, 13–23.

Innes, P.J., Tan, D.K.Y., Van Ogtrop, F., Amthor, J.S., 2015. Effects of high-temperature episodes on wheat yields in New South Wales, Australia. Agric. For. Meteorol. 208, 95–107.

Jones, C.A., Kiniry, J.R., 1986. CERES-Maize; a simulation model of maize growth and development.

Jones, J.W., Hoogenboom, G., Porter, C., Boote, K., Batchelor, W., Hunt, L., Ritchie, J., 2003. The DSSAT cropping system model European Journal of Agronomy.

Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Wang, E., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., McCown, R.L., Freebairn, D.M., Smith, C.J., 2003. An overview of the crop model. Eur. J. Agron. 18, 267–288. https://doi.org/10.1016/S0223-5234(03)00100-4

Keating, B.A., Gaydon, D., Huth, N.I., Probert, M.E., Verburg, K., Smith, C.J., Bond, W., 2002. Use of modelling to explore the water balance of dryland farming systems in the Murray-Darling Basin, Australia. Eur. J. Agron. 18, 159–169.

Keating, B.S., Meinke, H., Probert, M.E., Huth, N.I., Hills, I.G., others, 2001. NWheat: documentation and performance of a wheat module for APSIM.

Kundzewicz, Z.W., Kanae, S., Seneviratne, S.I., Handmer, J., Nicholls, N., Peduzzi, P., Mechler, R., Bouwer, L.M., Arnell, N., Mach, K., others, 2014. Flood risk and climate change: global and regional perspectives. Hydrol. Sci. J. 59, 1–28.

Lawes, R.A., Oliver, Y.M., Robertson, M.J., 2009. Integrating the effects of climate and plant available soil water holding capacity on wheat yield. F. Crop. Res. 113, 297–305.

Li, J., Wang, E., Wang, Y., Xing, H., Wang, D., Wang, L., Gao, C., 2016. Reducing greenhouse gas emissions from a wheat--maize rotation system while still maintaining productivity. Agric. Syst. 145, 90–98.

Li, K., Yang, X., Liu, Z., Zhang, T., Lu, S., Liu, Y., 2014. Low yield gap of winter wheat in the North China Plain. Eur. J. Agron. 59, 1–12.

Lilley, J.M., Kirkegaard, J.A., 2008. Seasonal variation in the value of subsoil water to wheat: simulation studies in southern New South Wales. Aust. J. Agric. Res. 58, 1115–1128.

Lilley, J.M., Kirkegaard, J.A., Robertson, M.J., Probert, M.E., Angus, J.F., Howe, G., 2003. Simulating crop and soil processes in crop sequences in southern NSW, in: Proceedings of the 11th Australian Agronomy Conference'. Geelong.

Littleboy, M., Silburn, D.M., Freebairn, D.M., Woodruff, D.R., Hammer, G.L., Leslie, J.K., 1992. Impact of soil erosion on production in cropping systems. I. Development and validation of a simulation model. Soil Res. 30, 757–774.

Liu, B., Asseng, S., Liu, L., Tang, L., Cao, W., Zhu, Y., 2016a. Testing the responses of four wheat crop models to heat stress at anthesis and grain filling. Glob. Chang. Biol. 22, 1890–1903.

Liu, B., Asseng, S., Müller, C., Ewert, F., Elliott, J., Lobell, D.B., Martre, P., Ruane, A.C., Wallach, D., Jones, J.W., Rosenzweig, C., Aggarwal, P.K., Alderman, P.D., Anothai, J., Basso, B., Biernath, C., Cammarano, D., Challinor, A., Deryng, D., De Sanctis, G., Doltra, J., Fereres, E., Folberth, C., Garcia-Vila, M., Gayler, S., Hoogenboom, G., Hunt, L.A., Izaurralde, R.C., Jabloun, M., Jones, C.D., Kersebaum, K.C., Kimball, B.A., Koehler, A.K., Kumar, S.N., Nendel, C., O'Leary, G.J., Olesen, J.E., Ottman, M.J., Palosuo, T., Prasad, P.V.V., Priesack, E., Pugh, T.A.M., Reynolds, M., Rezaei, E.E., Rötter, R.P., Schmid, E., Semenov, M.A., Shcherbak, I., Stehfest, E., Stöckle, C.O., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P., Waha, K., Wall, G.W., Wang, E., White, J.W., Wolf, J., Zhao, Z., Zhu, Y., 2016b. Similar estimates of temperature impacts on global wheat yield by three independent methods. Nat. Clim. Chang. 6, 1130–1136. https://doi.org/10.1038/nclimate3115

Lobell, D.B., Sibley, A., Ivan Ortiz-Monasterio, J., 2012. Extreme heat effects on wheat senescence in India. Nat. Clim. Chang. 2, 186–189. https://doi.org/10.1038/nclimate1356

Lobell, Sibley, A., Ortiz-Monasterio, J.I., 2012. Extreme heat effects on wheat senescence in India. Nat. Clim. Chang. 2, 186–189. https://doi.org/10.1038/nclimate1356

Luck, J., Spackman, M., Freeman, A., Trębicki, P., Griffiths, W., Finlay, K., Chakraborty, S., 2011. Climate change and diseases of food crops. Plant Pathol. 60, 113–121.

Ludwig, F., Asseng, S., 2006. Climate change impacts on wheat production in a Mediterranean environment in Western Australia. Agric. Syst. 90, 159–179.

Luo, Q., Bellotti, W., Williams, M., Bryan, B., 2005. Potential impact of climate change on wheat yield in South Australia. Agric. For. Meteorol. 132, 273–285.

Maiorano, A., Martre, P., Asseng, S., Ewert, F., Müller, C., Rötter, R.P., Ruane, A.C., Semenov, M.A., Wallach, D., Wang, E., Alderman, P.D., Kassie, B.T., Biernath, C., Basso, B., Cammarano, D., Challinor, A.J., Doltra, J., Dumont, B., Rezaei, E.E., Gayler, S., Kersebaum, K.C., Kimball, B.A., Koehler, A.K., Liu, B., O'Leary, G.J., Olesen, J.E., Ottman, M.J., Priesack, E., Reynolds, M., Stratonovitch, P., Streck, T., Thorburn, P.J., Waha, K., Wall, G.W., White, J.W., Zhao, Z., Zhu, Y., 2017. Crop model improvement reduces the uncertainty of the response to temperature of multi-model ensembles. F. Crop. Res. 202, 5–20. https://doi.org/10.1016/j.fcr.2016.05.001

Manschadi, A.M., Christopher, J., deVoil, P., Hammer, G.L., 2006. The role of root architectural traits in adaptation of wheat to water-limited environments. Funct. plant Biol. 33, 823–837.

Martre, P., Quilot-Turion, B., Luquet, D., Memmah, M.-M.O.-S., Chenu, K., Debaeke, P., 2015a. Model-assisted phenotyping and ideotype design, in: Crop Physiology. Elsevier, pp. 349–373.

Martre, P., Wallach, D., Asseng, S., Ewert, F., Jones, J.W., Rötter, R.P., Boote, K.J., Ruane, A.C., Thorburn, P.J., Cammarano, D., others, 2015b. Multimodel ensembles of wheat growth: many models are better than one. Glob. Chang. Biol. 21, 911–925.

McCown, R.L., Hammer, G.L., Hargreaves, J.N.G., Holzworth, D., Huth, N.I., 1995. APSIM: an agricultural production system simulation model for operational research. Math. Comput. Simul. 39, 225–231. https://doi.org/10.1016/0378-4754(95)00063-2

McCown, R.L., Hammer, G.L., Hargreaves, J.N.G., Holzworth, D.P., Freebairn, D.M., 1996. APSIM- A novel software system for model development.pdf. Agric. Syst. 50, 255–271.

Meehl, G.A., Zwiers, F., Evans, J., Knutson, T., Mearns, L., Whetton, P., 2000. Trends in extreme weather and climate events: issues related to modeling extremes in projections of future climate change. Bull. Am. Meteorol. Soc. 81, 427–436.

Mielenz, H., Thorburn, P.J., Scheer, C., Migliorati, M.D.A., Grace, P.R., Bell, M.J., 2016. Opportunities for mitigating nitrous oxide emissions in subtropical cereal and fiber cropping systems: A simulation study. Agric. Ecosyst. Environ. 218, 11–27.

Moeller, C., Pala, M., Manschadi, A.M., Meinke, H., Sauerborn, J., 2007. Assessing the sustainability of wheat-based cropping systems using APSIM: model parameterisation and evaluation. Aust. J. Agric. Res. 58, 75–86.

Mohanty, M., Probert, M.E., Reddy, K.S., Dalal, R.C., Mishra, A.K., Subba Rao, A., Singh, M., Menzies, N.W., 2012. Simulating soybean-wheat cropping system: APSIM model parameterization and validation. Agric. Ecosyst. Environ. 152, 68–78. https://doi.org/10.1016/j.agee.2012.02.013

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2011. Soil and water assessment tool theoretical documentation version 2009.

O'Leary, G.J., Christy, B., Nuttall, J., Huth, N., Cammarano, D., Stöckle, C., Basso, B., Shcherbak, I., Fitzgerald, G., Luo, Q., others, 2015. Response of wheat growth, grain yield and water use to elevated CO 2 under a Free-Air CO 2 Enrichment (FACE) experiment and modelling in a semi-arid environment. Glob. Chang. Biol. 21, 2670–2686.

O'Leary, G.J., Li Liu, D., Ma, Y., Li, F.Y., McCaskill, M., Conyers, M., Dalal, R., Reeves, S., Page, K., Dang, Y.P., others, 2016. Modelling soil organic carbon 1. Performance of APSIM crop and pasture modules against long-term experimental data. Geoderma 264, 227–237.

Oliver, Y., Wong, M., Robertson, M., Wittwer, K., others, 2006. PAWC determines spatial variability in grain yield and nitrogen requirement by interacting with rainfall on northern WA sandplain, in: Proceedings of the 13th Australian Agronomy Conference. pp. 10–14.

Oliver, Y.M., Robertson, M.J., 2009. Quantifying the benefits of accounting for yield potential in spatially and seasonally responsive nutrient management in a Mediterranean climate. Soil Res. 47, 114–126.

Oliver, Y.M., Robertson, M.J., Stone, P.J., Whitbread, A., 2009. Improving estimates of water-limited yield of wheat by accounting for soil type and within-season rainfall. Crop pasture Sci. 60, 1137–1146.

Paydar, Z., Huth, N., Ringrose-Voase, A., Young, R., Bernardi, T., Keating, B., Cresswell, H., 2005. Deep drainage and land use systems. Model verification and systems comparison. Aust. J. Agric. Res. 56, 995–1007.

Peake, A.S., Huth, N.I., Carberry, P.S., Raine, S.R., Smith, R.J., 2014. Quantifying potential yield and lodging-related yield gaps for irrigated spring wheat in sub-tropical Australia. F. Crop. Res. 158, 1–14.

Perkins, S.E., Alexander, L. V, Nairn, J.R., 2012. Increasing frequency, intensity and duration of observed global heatwaves and warm spells. Geophys. Res. Lett. 39.

Phelan, D.C., Harrison, M.T., McLean, G., Cox, H., Pembleton, K.G., Dean, G.J., Parsons, D., do Amaral Richter, M.E., Pengilley, G., Hinton, S.J., others, 2018. Advancing a farmer decision support tool for agronomic decisions on rainfed and irrigated wheat cropping in Tasmania. Agric. Syst. 167, 113–124.

Probert, M.E., Carberry, P.S., McCown, R.L., Turpin, J.E., 1998. Simulation of legume-cereal systems using APSIM. Aust. J. Agric. Res. 49, 317–328.

Probert, M.E., Dimes, J.P., 2004. Modelling release of nutrients from organic resources using APSIM, in: ACIAR PROCEEDINGS. pp. 25–31.

Probert, M.E., Keating, B.A., Thompson, J.P., Parton, W.J., 1995. Modelling water, nitrogen, and crop yield for a long-term fallow management experiment. Aust. J. Exp. Agric. 35, 941–950.

Raes, D., Steduto, P., Hsiao, T.C., Fereres, E., 2009. AquaCrop—the FAO crop model to simulate yield response to water: II. Main algorithms and software description. Agron. J. 101, 438–447.

Rigby, J.R., Porporato, A., 2008. Spring frost risk in a changing climate. Geophys. Res. Lett. 35.

Roberts, M.J., Braun, N.O., Sinclair, T.R., Lobell, D.B., Schlenker, W., 2017. Comparing and combining process-based crop models and statistical models with some implications for climate change. Environ. Res. Lett. 12, 95010.

Robertson, M.J., Carberry, P.S., Huth, N.I., Turpin, J.E., Probert, M.E., Poulton, P.L., Bell, M., Wright, G.C., Yeates, S.J., Brinsmead, R.B., 2002. Simulation of growth and development of diverse legume species in APSIM. Aust. J. Agric. Res. 53, 429–446.

Robertson, M.J., Lilley, J.M., 2016. Simulation of growth, development and yield of canola (Brassica napus) in APSIM. Crop Pasture Sci. 67, 332–344.

Rötter, R.P., Tao, F., Höhn, J.G., Palosuo, T., 2015. Use of crop simulation modelling to aid ideotype design of future cereal cultivars. J. Exp. Bot. 66, 3463–3476.

Sadras, V., Baldock, J., Roget, D., Rodriguez, D., 2003. Measuring and modelling yield and water budget components of wheat crops in coarse-textured soils with chemical constraints. F. Crop. Res. 84, 241–260.

Shamudzarira, Z., Robertson, M.J., 2002. Simulating response of maize to nitrogen fertilizer in semi-arid Zimbabwe. Exp. Agric. 38, 79–96.

Shiferaw, B., Smale, M., Braun, H.-J., Duveiller, E., Reynolds, M., Muricho, G., 2013. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. Food Secur. 5, 291–317.

Shroyer, J.P., Mikesell, M.E., Paulsen, G.M., 1995. Spring freeze injury to Kansas wheat. Cooperative Extension Service, Kansas State University.

Soltani, A., Sinclair, T.R., 2015. A comparison of four wheat models with respect to robustness and transparency: Simulation in a temperate, sub-humid environment. F. Crop. Res. 175, 37–46.

Steduto, P., Hsiao, T.C., Raes, D., Fereres, E., 2009. AquaCrop—The FAO crop model to simulate yield response to water: I. Concepts and underlying principles. Agron. J. 101, 426–437.

Stone, P.J., Nicolas, M.E., 1995. A survey of the effects of high temperature during grain filling on yield and quality of 75 wheat cultivars. Aust. J. Agric. Res. 46, 475–492.

Sun, H., Zhang, X., Wang, E., Chen, S., Shao, L., 2015. Quantifying the impact of irrigation on groundwater reserve and crop production--a case study in the North China Plain. Eur. J. Agron. 70, 48–56.

Tao, F., Rötter, R.P., Palosuo, T., Gregorio Hernández D\'\iaz-Ambrona, C., M\'\inguez, M.I., Semenov, M.A., Kersebaum, K.C., Nendel, C., Specka, X., Hoffmann, H., others, 2018. Contribution of crop model structure, parameters and climate projections to uncertainty in climate change impact assessments. Glob. Chang. Biol. 24, 1291–1307.

Thorburn, P.J., Probert, M.E., Robertson, F.A., 2001. Modelling decomposition of sugar cane surface residues with APSIM--Residue. F. Crop. Res. 70, 223–232.

Trenberth, K.E., 2011. Changes in precipitation with climate change. Clim. Res. 47, 123–138.

Van Dam, J.C., Huygen, J., Wesseling, J.G., Feddes, R.A., Kabat, P., Van Walsum, P.E. V, Groenendijk, P., Van Diepen, C.A., 1997. Theory of SWAP version 2.0 Simulation of water flow, solute transport and plant growth in the soil-water-atmosphere-plant environment.

Van Diepen, C.A. van, Wolf, J., Van Keulen, H., Rappoldt, C., 1989. WOFOST: a simulation model of crop production. Soil use Manag. 5, 16–24.

van Ittersum, M.K., Leffelaar, P.A., Van Keulen, H., Kropff, M.J., Bastiaans, L., Goudriaan, J., 2003. On approaches and applications of the Wageningen crop models. Eur. J. Agron. 18, 201–234.

Van Oort, P.A.J., Wang, G., Vos, J., Meinke, H., Li, B.G., Huang, J.K., van der Werf, W., 2016. Towards groundwater neutral cropping systems in the Alluvial Fans of the North China Plain. Agric. Water Manag. 165, 131–140.

Van Wijk, M.T., Rufino, M.C., Enahoro, D., Parsons, D., Silvestri, S., Valdivia, R.O., Herrero, M., 2014. Farm household models to analyse food security in a changing climate: A review. Glob. Food Sec. 3, 77–84.

Vrugt, J.A., Ter Braak, C.J.F., Clark, M.P., Hyman, J.M., Robinson, B.A., 2008. Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. Water Resour. Res. 44.

Wang, E., Chen, C., Yu, Q., 2009. Modeling the response of wheat and maize productivity to climate variability and irrigation in the North China Plain, in: 18th World IMACS/MODSIM Congress. pp. 2742–2748.

Wang, E., Martre, P., Zhao, Z., Ewert, F., Maiorano, A., Rötter, R.P., Kimball, B.A., Ottman, M.J., Wall, G.W., White, J.W., others, 2017. The uncertainty of crop yield projections is reduced by improved temperature response functions. Nat. Plants 3, 17102.

Wang, E., Van Oosterom, E., Meinke, H., Asseng, S., Robertson, M.J., Huth, N., Keating, B., Probert, M., 2003. The new APSIM-Wheat Model—performance and future improvements, in: Proceedings of the 11th Australian Agronomy Conference.

Wang, G.-C., Wang, E.-L., Yao, H., Xu, J.-J., 2014. Soil carbon sequestration potential as affected by management practices in northern China: a simulation study. Pedosphere 24, 529–543.

Wang, J., Wang, E., Feng, L., Yin, H., Yu, W., 2013. Phenological trends of winter wheat in response to varietal and temperature changes in the North China Plain. F. Crop. Res. 144, 135–144.

Williams, J.R., Jones, C.A., Kiniry, J.R., Spanel, D.A., 1989. The EPIC crop growth model. Trans. ASAE 32, 497–511.

Wong, M.T.F., Asseng, S., 2006. Determining the causes of spatial and temporal variability of wheat yields at sub-field scale using a new method of upscaling a crop model. Plant Soil 283, 203–215.

Xiao, D., Tao, F., 2014. Contributions of cultivars, management and climate change to winter wheat yield in the North China Plain in the past three decades. Eur. J. Agron. 52, 112–122.

Yan, Z., Zhang, X., Rashid, M.A., Li, H., Jing, H., Hochman, Z., 2020. Assessment of the sustainability of different cropping systems under three irrigation strategies in the North China Plain under climate change. Agric. Syst. 178, 102745.

Yang, Yanmin, Li Liu, D., Anwar, M.R., Zuo, H., Yang, Yonghui, 2014. Impact of future climate change on wheat production in relation to plant-available water capacity in a semiaridenvironment. Theor. Appl. Climatol. 115, 391–410.

Yunusa, I.A.M., Bellotti, W.D., Moore, A.D., Probert, M.E., Baldock, J.A., Miyan, S.M., 2004. An exploratory evaluation of APSIM to simulate growth and yield processes for winter cereals in rotation systems in South Australia. Aust. J. Exp. Agric. 44, 787–800.

Zadoks, J.C., Chang, T.T., Konzak, C.F., 1974. A decimal code for the growth stages of cereals. Weed Res. 14, 415–421.

Zeleke, K.T., Nendel, C., 2016. Analysis of options for increasing wheat (Triticum aestivum L.) yield in south-eastern Australia: The role of irrigation, cultivar choice and time of sowing. Agric. Water Manag. 166, 139–148.

Zeppel, M.J.B., Wilks, J. V, Lewis, J.D., 2014. Impacts of extreme precipitation and seasonal changes in precipitation on plants. Biogeosciences 11, 3083–3093.

Zhang, H., Franssen, H.J.H., Han, X., Vrugt, J.A., Vereecken, H., 2017. Joint state and parameter estimation of two land surface models using the ensemble Kalman filter and particle filter. Hydrol. Earth Syst. Sci. 21, 4927–4958. https://doi.org/10.5194/hess-21-4927-2017

Zhang, Y., Feng, L., Wang, E., Wang, J., Li, B., 2012. Evaluation of the APSIM-Wheat model in terms of different cultivars, management regimes and environmental conditions. Can. J. Plant Sci. 92, 937–949.

Zhang, Y., Feng, L.P., Wang, J., Wang, E.L., Xu, Y.L., 2013. Using APSIM to explore wheat yield response to climate change in the North China Plain: the predicted adaptation of wheat cultivar types to vernalization. J. Agric. Sci. 151, 836–848.

Zhao, J., Pu, F., Li, Y., Xu, J., Li, N., Zhang, Y., Guo, J., Pan, Z., 2017. Assessing the combined effects of climatic factors on spring wheat phenophase and grain yield in Inner Mongolia, China. PLoS One 12.

Zhao, Z., Qin, X., Wang, E., Carberry, P., Zhang, Y., Zhou, S., Zhang, X., Hu, C., Wang, Z., 2015. Modelling to increase the eco-efficiency of a wheat--maize double cropping system. Agric. Ecosyst. Environ. 210, 36–46.

Zhao, Z., Wang, E., Wang, Z., Zang, H., Liu, Y., Angus, J.F., 2014a. A reappraisal of the critical nitrogen concentration of wheat and its implications on crop modeling. F. Crop. Res. 164, 65–73.

Zhao, Z., Wang, E., Xue, L., Wu, Y., Zang, H., Qin, X., Zhang, J., Wang, Z., 2014b. Accuracy of root modelling and its impact on simulated wheat yield and carbon cycling in soil. F. Crop. Res. 165, 99–110.

Zheng, B., Chapman, S.C., Christopher, J.T., Frederiks, T.M., Chenu, K., 2015. Frost trends and their estimated impact on yield in the Australian wheatbelt. J. Exp. Bot. 66, 3611–3623.

Zheng, B., Chenu, K., Doherty, A., Chapman, S., 2014. The APSIM-wheat module (7.5 R3008). Agric. Prod. Syst. Simulator Initiat.

Zhou, S., Zhang, Y., Williams, A.P., Gentine, P., 2019. Projected increases in intensity, frequency, and terrestrial carbon costs of compound drought and aridity events. Sci. Adv. 5, eaau5740.

Click here to access/download
Supplementary Material
Table_S1_Calibrated Cultivar parameters of APSIM-Wheat.docx

Click here to access/download
Supplementary Material
Table_S2_Detailed Information of validation datasets.docx

Conflict of Interest of *Performance of a wheat yield prediction model and factors influencing the performance: A review and meta-analysis*

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