Agricultural and Forest Meteorology 288–289 (2020) 107995

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

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

Comparison of wheat simulation models for impacts of extreme temperature stress on grain quality

Raheel Osman, Yan Zhu, Wei Ma, Dongzheng Zhang, Zhifeng Ding, Leilei Liu, Liang Tang, Bing Liu*, Weixing Cao*

National Engineering and Technology Center for Information Agriculture, Engineering Research Center for Smart Agriculture, Ministry of Education, Key Laboratory for Crop System Analysis and Decision Making, Ministry of Agricultureand Rural Affairs, Jiangsu Key Laboratory for Information Agriculture, Jiangsu Collaborative Innovation Center for Modern Crop Production, Nanjing Agricultural University, Nanjing, Jiangsu 210095, China

ARTICLE INFO

Keywords:
Crop models
Grain quality
High-temperature stress
Low-temperature stress
Model evaluation

ABSTRACT

Shifting temperature patterns on global and regional scales accredited to climate change will bring more low-temperature stress (LTS) and high-temperature stress (HTS) events to further deteriorate wheat yield and quality. Crop models can serve as a beneficial platform for quantifying the impact of both high and low-temperature events on grain yield (GY) and grain quality. Wheat grain protein concentration (GPC) and grain protein yield (GPY), as two important measurements of wheat quality for nutrition value, is often ignored in crop modeling efforts to improve grain yield under climate change. This study was undertaken for comprehensive comparison of four broadly used wheat simulation models (DSSAT-CERES-Wheat, DSSAT-Nwheat, WheatGrow, and APSIM-Wheat) in quantifying and simulating the responses of wheat grain quality (GPC and GPY) under LTS and HTS at critical growth stages, and to identify gaps in simulating wheat grain protein concentration and protein yield for crop model improvement. Four-year environment-controlled phytotron experiments were conducted with two wheat varieties under LTS (at joining and booting stages) and HTS (at anthesis, grain filling, and combined stress at anthesis and grain filling stages). For per unit increase in cold degree days (CDD, degree days below 2 °C) at jointing and booting stages and heat degree days (HDD, degree days over 30 °C) at anthesis, grain filling and combined stress at anthesis and grain filling stages, GPC was increased by 0.2% to 0.4% and 1.1% to 1.6%, while GPY was decreased by 2.1% to 4.5% and 0.3% to 1.7%, respectively. Most of the crop models tended to reproduce some HTS impacts better during grain filling than at anthesis, but all the tested models call for improvements in simulating LTS at different stages, especially for GPY. Our results indicated the need of incorporating response functions of extreme temperature stresses into grain quality models to adapt to future climate scenarios.

1. Introduction

Temperature anomalies as a consequence of global climate change have already produced significant negative effects on agricultural production and quality (Kawasaki and Uchida, 2016). Shifting temperature patterns at global and regional scales will bring more low-temperature stress (LTS) events due to advanced phenology and high-temperature stress (HTS) events due to increasing temperature variability, to further deteriorate crop production. Wheat is the most important staple food crop in the world, and is sensitive to extreme temperature stress (Porter and Gawith, 1999), which has negative impacts on wheat growth and development, and deteriorates grain yield (GY) and quality (Dias and Lidon, 2009; Kawasaki and Uchida, 2016; Liu et al., 2016a; Porter and Gawith, 1999). Due to climate change, high-temperature events have been increased during the growing season of wheat crop (Asseng et al., 2011), and increasing HTS events were also projected under future climate scenarios (Teixeira et al., 2013). In addition, increasing growing season temperatures for major wheat producing areas (e.g. China and the Indo-Gangetic Plains of India) hastens wheat growth and development. This leads to the significant progression of sensitive growth stages to low-temperature, thus increases the probability of frost or cold damage to wheat (Gu et al., 2008; Lobell et al., 2012; Zheng et al., 2015). As a result, serious losses in wheat production due to extreme temperature events have been observed in Australia (Barlow et al., 2015; Crimp et al., 2016), Europe (Peings et al., 2013; Trnka et al., 2014), the United States (Augspurger, 2013; Gu, 2008),

* Corresponding authors.
E-mail addresses: bingliu@njau.edu.cn (B. Liu), caow@njau.edu.cn (W. Cao).

https://doi.org/10.1016/j.agrformet.2020.107995
Received 11 May 2019; Received in revised form 2 April 2020; Accepted 5 April 2020

and China (; Kodra et al., 2011; Liu et al., 2016a).

The composition and concentration of grain protein is an important quality measure that quantifies the end use, nutritional and rheological properties including loaf volume, extensibility, dough strength, and gluten (Nuttall et al., 2017). Further, grain protein is influenced by genotype, environmental conditions, and management practices (Asseng and Milroy, 2006; Spiertz et al., 2006). The rate and duration of protein deposition are determined mainly by supply factors external to the grain, e.g., climate conditions (Asseng and Milroy, 2006). However, extreme temperatures make it difficult to assure stable grain protein concentration (GPC) and the quality standards demanded by grain dealers (Kawasaki and Uchida, 2016). Therefore, to ensure quality standards, it is critical to maintain grain quality under climatic extremes (Nuttall et al., 2017).

LTS affects wheat throughout its life cycle from germination to maturity (Fuller et al., 2007; Porter and Gawith, 1999), and the reproductive stage is the most sensitive for GY formation (Frederiks et al., 2015). LTS is likely to damage the plant from double ridge stage and onward (Whaley et al., 2004), and usually occurs in the early stages of crop development (Frederiks et al., 2008; Fuller et al., 2007). Furthermore, recent studies show that due to climate change, vegetative stages (jointing and booting) become more prone to LTS particularly in China (Ji et al., 2017; Li et al., 2015). Most frequent damage of LTS occurring at booting was the loss of head in the top node of the stem (Frederiks et al., 2008), along with lower stem damage which causes lodging during grain filling period (Barlow et al., 2015), and finally result in the drastically reduced GY. Currently, there is no systematic study that focuses on the effects of LTS on GPC and grain protein yield (GPY).

Most previous studies concentrate on the impact of HTS on wheat GY and yield components, particularly at grain filling stage (Farooq et al., 2011; Lobell et al., 2012; Liu et al., 2014). A few studies focused on the impact of HTS on grain quality traits (DuPont and Altenbach, 2003; Kawasaki and Uchida, 2016; Wardlaw, 2002), specifically on GPC (Martre et al., 2003; Spiertz et al., 2006), at various critical growth stages. Wheat is very susceptible to HTS after heading (Porter and Gawith, 1999). Post-heading HTS, which is very common across global wheat-growing areas (Pradhan et al., 2012), results in accelerated crop development and limited grain growth during the grain filling period (Liu et al., 2016a; Lobell et al., 2012), and negatively affects grain quality (Liu et al., 2016a; Wardlaw, 2002). Generally, HTS increases wheat GPC, but reduces protein accumulation per grain significantly (Asseng and Milroy, 2006).

Process-based crop models, which account for the interactions among climate, crop, soil and management efforts, have become the most common tools for assessing the impact of climate change on crop productivity (Chenu et al., 2017). Recently, in the Agriculture Model Inter-comparison and Improvement Project (AgMIP), performance of 30 wheat crop models for predicting GY was evaluated under a wide range of temperature conditions (Asseng et al., 2015), and several wheat crop models were appeared to be less accurate under high-temperature. In addition, various other studies have elicited concerns regarding the simulation of extreme temperature effects with contemporary crop models to predict effects of global warming on future crop productivity (Kawasaki and Uchida, 2016; Sánchez et al., 2014). Craufurd et al. (2013) proposed the urgent need to design experiments in crop science to evaluate and improve the model performance under extreme temperatures. Liu et al. (2016a) have identified the shortcomings of four wheat models under heat stress at anthesis and grain filling periods. In addition, the study conducted by Siebert et al. (2014) indicated that canopy temperature should be considered to reduce uncertainty in assessing heat stress impacts on crop yield. Recently, model improvements for the impacts of HTS on crop growth and yield have been conducted, especially for HTS impacts on biomass, phenology, leaf senescence, and grain yield with air temperature (Liu et al., 2017, 2016b; Maiorano et al., 2017; Wang et al., 2017) and canopy temperature in wheat (Webber et al., 2017). However, most of these studies focused on wheat growth and yield parameters, not on the grain quality indices. Moreover, several studies only considered HTS and did not account for LTS (Kawasaki and Uchida, 2016). Furthermore, several calibration and validation of models in simulating grain quality were conducted only under various nitrogen and water conditions (Jamieson and Semenov, 2000; Martre et al., 2003), but not under extreme temperature conditions. Recently, some works with single model focused on the algorithm improvement for GPC simulation of durum wheat under normal temperature routine (Orlando et al., 2017) and simulations with multi-model ensembles were done by Yin et al. (2017) and Rötter et al. (2018). But these studies didn't consider the effect of extreme temperatures on grain quality formation.

Therefore, the objectives of this study were: (1) to conduct comprehensive comparison of four extensively used wheat models (DSSAT-CERES-Wheat, DSSAT-Nwheat, WheatGrow, and APSIM-Wheat) in simulating the responses of wheat grain quality (GPC and GPY) under LTS and HTS at critical growth stages; and (2) to identify gaps in simulating wheat GPC and GPY for future crop model improvement.

2. Materials and methods

2.1. Data sources

Four-year environment-controlled phytotron experiments were carried out at Rugao Experimental Station of Nanjing Agricultural University (120.33°E, 32.23°N) in Jiangsu Province of China during 2013–2017 growing seasons. The phytotrons (FYS-10, Yuheng Instrument Co. Ltd., Nanjing, China) with 3.4 m × 3.2 m × 2.8 m of length, width and height, were used to conduct extreme temperature experiments. The two of the most widely cultivated winter wheat varieties (cv. Yangmai16 and cv. Xumai30) in this region were planted in plastic pots. The height and diameter of plastic pot were approximately 0.3 × 0.28 m. All pots, with a plant density of 10 plants per pot, were kept under normal optimum conditions before and after the treatments of temperature stresses. Base fertilization at a rate of 183 kg ha^{-1} N, 102 kg ha^{-1} P_2O_5 and 183 kg ha^{-1} K_2O was applied to each pot before sowing, and an extra 183 kg ha^{-1} N per pot was top-dressed at jointing stage. The fertilization rate was about 30% higher than in the local fields to make sure no nutritional stress for wheat growth, because root growth depth in fields usually wasdeeper than the depth of pots used here. Irrigation and pesticide were applied according to local standards to ensure that HTS and LTS were the only limiting factors in the present study. Sowing dates of HTS experiments were November 2, November 3 for cv. Yangmai16 and November 5, and November 12 for cv. Xumai30 during growing season of 2014–2015 and 2015–2016, respectively. For LTS experiments, sowing dates of cv. Yangmai16 and cv. Xumai30 were November 4 and November 3 during growing season of 2014–2015 and 2015–2016, respectively.

Table 1 summarizes the treatmentsof LTS and HTS. In LTS experiments,there were four temperature levels (Tmin/Tmax: 6/16, −2/8, −4/6, and −6/4 °C), three stress durations (2, 4, and 6 days) and two treatment growth stages (jointing and booting). In HTS experiments, there were four temperature levels (Tmin/Tmax: 17/27, 21/31, 25/35, and 29/39), two stress durations (3 and 6 days), and three treatment stages (anthesis, grain filling and combined stress at anthesis and grain filling).

The pots were then divided into three blocks, each having 25 rows × 8 columns. These blocks were surrounded by additional pots to prevent side effects. Once the wheat plants reached to the anticipated growth stages, jointing (Zadok 31) and booting (Zadok 45) for LTS, and anthesis (when the first anther were observed at middle of ear (Zadok 61) and grain filling (10 days after the first anther at middle of ear or 10 days after Zadok 61) for HTS, pots were randomly selected from three blocks and shifted into phytotrons to expose them to various stress conditions.

Table 1
Summary of extreme temperature stress treatments in environment-controlled phytotron experiments.

Stress	Cultivar	Growing season	Starting time of treatment	Duration (days)	Temperature regime (T_{min}/T_{max}) (°C)
LTS	Yangmai16	2014–2015	Jointing‡, Booting†	D1(2), D2(4), D3(6)	T1(6/16), T2 (–2/8), T3 (–4/6), T4 (–6/4)
		2015–2016	Jointing‡, Booting†	D1(2), D2(4), D3(6)	T1(6/16), T2(−2/8), T3(−4/6), T4(−6/4)
	Xumai30	2014–2015	Jointing‡, Booting†	D1(2), D2(4), D3(6)	T1(6/16), T2(–2/8), T3(–4/6), T4(–6/4)
		2015–2016	Jointing‡, Booting†	D1(2), D2(4), D3(6)	T1(6/16), T2(−2/8), T3(−4/6), T4(−6/4)
HTS	Yangmai16	2015–2016	Anthesis*, Grain filling**, Combined stress at anthesis and grain filling	D1(3), D2(6)	T1(17/27), T2 (21/31), T3 (25/35), T4 (29/39)
		2016–2017	Anthesis*, Grain filling**, Combined stress at anthesis and grain filling	D1(3), D2(6)	T1(17/27), T2 (21/31), T3 (25/35), T4 (29/39)
	Xumai30	2015–2016	Anthesis*, Grain filling**, Combined stress at anthesis and grain filling	D1(3), D2(6)	T1(17/27), T2 (21/31), T3 (25/35), T4 (29/39)
		2016–2017	Anthesis*, Grain filling**, Combined stress at anthesis and grain filling	D1(3), D2(6)	T1(17/27), T2 (21/31), T3 (25/35), T4 (29/39)

‡ Jointing stage means the first node emerges above the soil line.
† Booting stage means when the flag leaf sheath is swollen.
* Anthesis means when the first anther was observed at the middle of ear.
** Grain filling means 10 days after the first anther observed.

The optimal temperature range for wheat growth during jointing and booting stages was 9.3–11.9 °C (Porter and Gawith, 1999). Therefore, T1 (6/16 °C) with an average temperature of 11 °C was selected as a control treatment. Additionally, the average lowest minimum and diurnal temperature from jointing and booting in the main wheat-growing regions of China was around −6 °C and 10 °C from 1981 to 2010 (Ji et al., 2017). Therefore, T4 (−6/4 °C) with an average temperature of −1 °C was selected as the lowest temperature level in this study. The average temperature during the growing season of 2014 to 2016 was 11 °C, 7 °C, 3 °C, and –1 °C, respectively, for LTS regimes of T1, T2, T3 and T4 (Table 1). In the case of HTS, 30 °C was selected as a threshold for winter wheat cultivars (Farooq et al., 2011; Liu et al., 2014). Hence, T1 (17/27 °C) with a maximum temperature of 27 °C considered as the optimal temperature for post-heading period in wheat and designed as the control treatment. T2 (21/31 °C), T3 (25/35 °C) and T4 (29/39 °C) with the maximum temperatures of 31 °C, 35 °C and 39 °C were designed as HTS treatments.

In order to simulate daily observed temperature fluctuations in the field environment, the relative humidity and air temperature were precisely controlled in the phytotrons by using a bubbling system and air conditioners. The fluctuations in diurnal temperature of the phytotrons during temperature treatments followed a similar pattern of temperature in the surrounding area (Fig. S1). The CO_2 concentration was maintained at the level of surrounding environment by installing two fans in each phytotron. The humidity inside the phytotrons was maintained at 70% which was consistent with surrounding environment during treatment periods. Temperature and relative humidity were recorded by using EM50 data loggers (Decagon Devices, Inc., Washington, USA) during the treatment period with an interval of 5 min. The light intensity inside phytotron was optimized by using a halogen lamp, and maintaining at 1380 $\mu molm^{-1}s^{-1}$ at noon with sunshine, and at 240 $\mu molm^{-1}s^{-1}$ under cloud cover. The positional effects of pots in phytotron were reduced by daily rotation of pots randomly. Pots after extreme temperature treatments were placed under ambient temperature until harvest. Meteorological data such as daily radiation, temperature, and rainfall during wheat growing seasons were collected using Dynamet-1 K (Dynamax Inc., Houston, TX, USA). The distance between the experimental station and the weather station was 20 m. Sampling (12 pots per treatment)of wheat grain was performed at physiological maturity to determine GY and quality parameters. The semi-micro-Kjeldahl method was used to determine the total nitrogen (N) concentration of grains (Williams, 1984). GPC was calculated by multiplying a conversion factor of 5.7 with the percentage of grain N.

2.2. Wheat simulation models

Simulation of wheat grain quality traits (GPC and GPY) under the treatments of extreme temperature stress was carried out by using four extensively used wheat simulation models. Further, the model comparison was carried out to evaluate the response of GPC and GPY under extreme temperature stresses. All of the models used in this study can simulate wheat development, growth and grain dry matter and grain N dynamics on a daily basis.

DSSAT-CERES-Wheat: The Decision Support System for Agrotechnology Transfer DSSAT-CERES-Wheat model (v4.6) (Ritchie, 1985) is one of the most broadly used crop models and its origins can go back to the modeling efforts of Joe Ritchie and his colleges in the 1970s. DSSAT-CERES-Wheat has been tested at a number of eco-sites in the USA, Europe, and Asia (Liu et al., 2016a; Palosuo et al., 2011; Timsina and Humphreys, 2006). In CERES-Wheat the growth and development are potentially influenced by temperature and daylight. Radiation use efficiency approach was used to model growth and development. The response to vernalization temperature makes it suitable for phenology simulations of spring and winter wheat (Ritchie, 1985). Several model evaluation studies that used the CERES-Wheat model generally focus on the cultivars, planting densities, sowing dates, mean temperatures, N application rates, and water conditions.

In CERES-Wheat, temperature thresholds for different growth processes were shown in Table S1. Impacts of daily mean temperature were considered in most of the growth processes (Asseng et al., 2015). CERES-Wheat applies independent functions for grain dry matter and grain N accumulation. The grain growth is divided into three phases: (i) lag phase, (ii) linear phase, and (iii) maturity phase. In the CERES-Wheat, the lag phase is delayed until 120 °C days, and dry grain matter initialized at 3.5 mg while grain N initialized at zero. The source and sink limitation approach were used to simulate the grain dry matter and N dynamics. The demands for grain dry matter and N were a function of the maximum daily kernel growth rate, the number of grain per unit area and temperature. Genetic variation for grain N accumulation was

not incorporated in this routine. GPC was calculated by multiplying grain N concentration by a factor of 5.7. The dry grain matter and grain N accumulation increase from daily mean temperature under 0–16 °C, then remains constant till 35 °C, and ceases at 45 °C (Alderman et al., 2013; Wang et al., 2017) (Fig.S2).

DSSAT-Nwheat: The DSSAT-Nwheat model employs the APSIM-Nwheat model after integrating it into the DSSAT platform (v.4.7). This model has also been tested worldwide under different agro-climatological and agronomic management conditions (Asseng and Milroy, 2006; Kassie et al., 2016; Liu et al., 2016a; O'Leary et al., 2015). It involves the modules of crop growth and development, soil, water, N, crop residues as well as their interaction within a soil-plant system,which is determined by daily weather dynamics (Keating et al., 2003). Light interception and radiation use efficiency approaches were used to calculate potential daily biomass production. Further to limit the potential growth and development, reduction factors were applied under sub-optimal water, N, and temperature (Asseng et al., 2011).

Table S1 and Fig. S2 represents the impacts of minimum, optimum and maximum temperatures on different physiological processes in DSSAT-Nwheat. When the maximum temperature is greater than 34 °C, the increased leaf senescence takes place to simulate HTS effects on wheat leaf senescence (Asseng et al., 2011). DSSAT-Nwheat followed the same routine of CERES-Wheat with three main modifications to simulate the grain dry matter and GPC. For example, firstly the grain N accumulation in grains was initialized with up to 3% at beginning of grain filling; secondly, the daily grain dry matter transfer was constrained so that accumulation was always at least at 1.23% N of the daily dry matter transfer; and thirdly upper boundary of grain N was constrained to 4% of daily dry matter transfer (Asseng et al., 2002; Asseng and Milroy, 2006). Like CERES-Wheat, DSSAT N-wheat did not consider the genotypic parameters for grain N accumulation (Asseng et al., 2002; Asseng and Milroy, 2006). The accumulation rate of grain dry matter and N increases from base temperature (0 °C) to optimal temperature (26 °C and 25 °C for grain dry matter and N, respectively), then keep at the constant values (Wang et al., 2017) (Fig. S2).

WheatGrow: The WheatGrow model comprises five sub-models including apical and phenological development, photosynthesis and biomass accumulation, biomass partitioning and organ growth, yield and quality formation, and soil nitrogen dynamics and water balance. Evaluation studies with WheatGrow model in different ecological regions of China indicated that it executed well in simulating wheat growth and development (Liu et al., 2016a; Lv et al., 2013).

In WheatGrow model, hourly temperature calculated from daily maximum and minimum temperatures were used in quantifying temperature effects in most of the growth processes. However, WheatGrow (v.3.1) contains several new high-temperature functions to improve the predictions of wheat phenology, photosynthesis and biomass accumulation, biomass partitioning, and GY under heat stress (Liu et al., 2016a; 2017). In these functions, maximum temperature and heat degree days calculated from hourly temperature were used to quantify heat stress effects. More description of these functions can be found in the supplementary methods.WheatGrow model applies a separate module for the rate of grain N accumulation on a single grain basis and also accounts for the genotypic parameter for the maximum rate of grain N accumulation (mg N grain^{-1} day^{-1}). Curvilinear temperature response from 0 to 1 affects the daily rate of grain N accumulation. The daily mean temperature was used to quantify the effects of temperature on the rate of grain N accumulation, and 24.2 °C was considered as optimal temperature (Pan et al., 2006).

APSIM-Wheat: The Agricultural Production Systems Simulator (APSIM) for wheat (v7.9) recognized as a widely-used crop model with highly-recognized capabilities in simulating many processes in agricultural systems. APSIM contains a suite of modules that enable the simulation of systems for a diverse range of plant, soil, climate, management, and their interactions. Radiation interception and radiation use efficiency were used to calculate the potential biomass and growth and were limited by soil water deficiency (Keating et al., 2003). APSIM contains modules related to crop growth and development, soil, water, N, and crop residues as well as their interaction (Keating et al., 2003). It has been widely evaluated under diverse conditions such as CO_2 concentrations, N, water, temperature, and planting dates (Liu et al., 2016a; Lobell et al., 2012; O'Leary et al., 2015).

In APSIM-Wheat, most of the physiological processes are affected by daily mean air temperature except vernalization which is affected by canopy, minimum, and maximum temperatures. When the maximum temperature increases beyond 34 °C, the accelerated leaf senescence occurs to quantify HTS effects (Wang et al., 2017). APSIM-Wheat applied different routines for accumulations of grain dry matter and grain N which start just after anthesis and is the function of potential grain filling rate (g grain^{-1}°Cd^{-1}), grain number, N limiting factor of grain filling and daily mean temperature. The accumulation rates of grain dry matter and N increase with temperature within the range of 0–26 °C and of 0–25 °C, respectively, same with DSSAT-Nwheat. Further, APSIM-Wheat considers the genotypic parameter for the potential rate of grain N accumulation. The default accumulation rate of grain N is 0.000055 g grain^{-1} day^{-1} (Zheng et al., 2014). Table S1 represents different temperature responses used to simulate the physiological processes in APSIM-Wheat.

2.3. Model calibration and validation

Four wheat simulation models were calibrated and evaluated using independent experimental datasets. Model calibration was carried out with the dataset of one growing season, for example, the dataset of 2014–2015 for LTS and dataset of 2015–2016 for HTS. While the rest of the dataset was used to evaluate the model performance. Only cultivar parameters were adjusted to fit the simulations to the experimental dataset. Manual adjustment of cultivar parameters was made to fit the observed and simulated values, including phenology (heading, anthesis, and maturity), aboveground biomass at maturity, GY and GPC. As suggested in previous studies Liu et al., 2016a; Palosuo et al., 2011; Rötter et al., 2012), a better picture of model performance can be achieved by the combined use of statistical indicators, especially when comparing model performance. Therefore, three statistical indices were employed for comparing and assessing model performance, including normalized root mean square error (NRMSE), mean bias error (MBE), and model efficiency coefficient (EF), which can be calculated as Eqs. (1) to ((3).

$$NRMSE = \frac{\sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{sim,i})^2}{n}}}{\bar{X}_{obs}} \quad (1)$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(X_{sim,i} - X_{obs,i}) \quad (2)$$

$$EF = 1 - \frac{\sum_{i=1}^{n}(X_{obs,i} - X_{sim,i})^2}{\sum_{i=1}^{n}(X_{obs,i} - \bar{X}_{obs})^2} \quad (3)$$

Where $X_{obs,i}$, $X_{sim,i}$ were the observed and simulated values for the ith dataset,$\bar{X}_{obs}$ was the mean values of observed datasets.

NRMSE evaluates the average relative deviation between observed and simulated values in percentage. The NRMSE value >10% shows excellent model performance, NRMSE between 10 and 30% presents the satisfactory model predictions, and >30% depicts poor performance. MBE is the average prediction error representing the systematic error of a simulation model to under or over the forecast. EF is the summation of absolute squared differences among the simulated and observed values normalized by the variance of the observed values. The range of EF lies between -∞ and 1. The above mentioned statistical indices were the same as Rötter et al. (2011) and Liu et al. (2017).

Observed and simulated responses of grain protein concentration

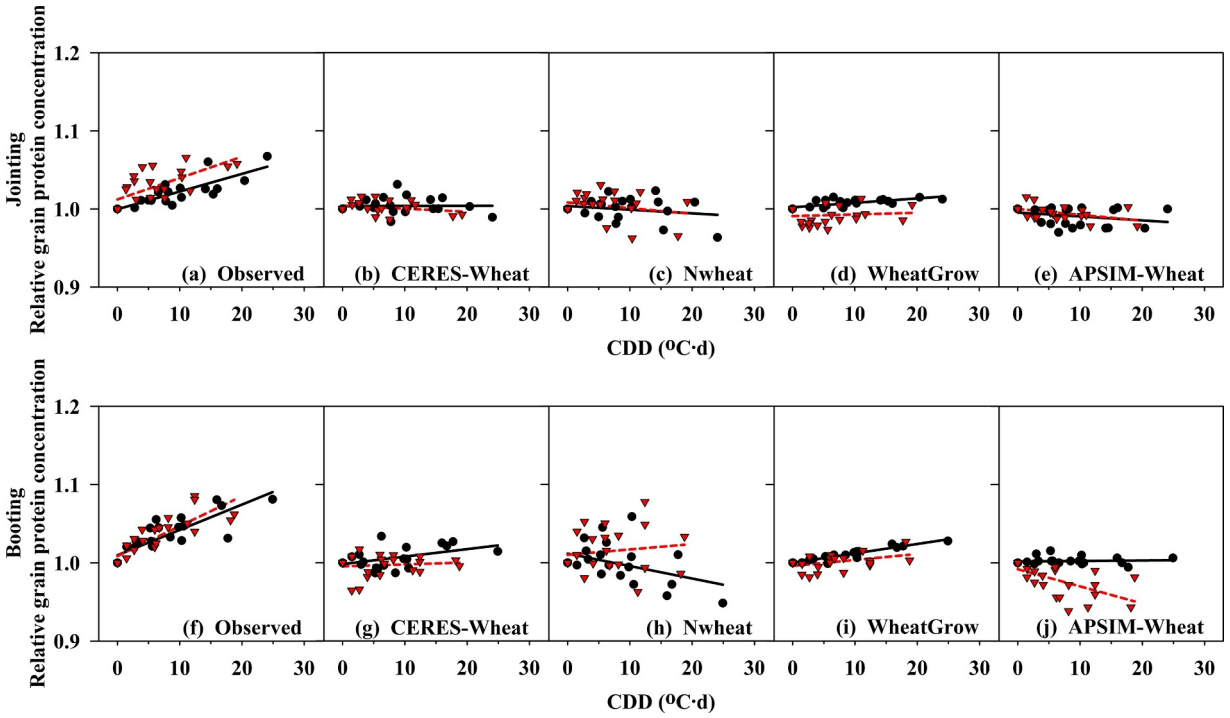

Fig. 1. Relative response of observed and simulated grain protein concentration to low temperature stress at jointing (a–e) and booting (f–j) from four wheat models for two winter wheat cultivars in environment-controlled phytotron experiments. Observed (a and f) and simulated with CERES-Wheat (b and g), Nwheat (c and h), WheatGrow (d and i), and APSIM-Wheat (e and j). Black circle and straight line indicate cv. Yangmai16; red triangle and dash line indicate cv. Xumai30.

(GPC,%) and grain protein yield (GPY, kg/ha) were tested individually under LTS at jointing and booting and under HTS at anthesis, grain filling and combined stress at anthesis and grain filling. The present study employed a simple linear regression to compare the observed and simulated responses of GPC and GPY to extreme temperatures. The regression slopes between observed or simulated wheat quality variables and extreme temperatures were considered as the observed or simulated responses of GPC and GPY to extreme temperature stresses. Relative values of GPC and GPY were used to calculate the relative responses of quality variables (GPC and GPY) to extreme temperatures, due to the variation in observed absolute values of GPC and GPY among different years and cultivars. The relative values of GPC and GPY were determined as the ratio between the absolute values under different treatments and the corresponding values from the control treatment (T1 treatment) for the same treatment stage and cultivar. ANOVA was applied to each regression to analyze the significance. Furthermore, Fig. S8-S13 provided a 1 to 1 comparison between observed and simulated wheat grain yield and quality indices from four models under LTS and HTS.

Cold degree days (CDD,°C d) and heat degree days (HDD,°C d) was used to quantify LTS and HTS in the regressions, respectively. CDD and HDD was the accumulated daily cold (CD*i*) and heat degree days (HD*i*) on the *i*th day during the stress treatment period, respectively. CDD and HDD reflect the intensity and duration of LTS and HTS across the treatments with different temperature levels and stress durations. No extreme temperature stress event was observed outside the extreme temperature stress treatments. The CDD was calculated as the following equations:

$$CDD = \sum_{i=1}^{m} CD_i \tag{4}$$

$$CD_i = \frac{1}{24}\sum_{h=1}^{24} CD_h \tag{5}$$

$$CD_{\mathrm{h}} = \begin{cases} 0 & T_h > T_t \\ T_t - T_h & T_{\mathrm{h}} \le T_t \end{cases} \tag{6}$$

Where CD_h indicates the hourly cold degree days on the*i*th day. T_t is the LTS threshold, which was set to 2 °C(Crimp et al., 2015; 2016; Potithep and Yasuoka, 2011); T_h (°C) corresponds to hourly temperature, which was recorded by EM50 data loggers. *m* in Eq. (4) is the cold stress duration.

HDD was calculated as follows:

$$HDD = \sum_{i=1}^{m} HD_i \tag{7}$$

$$HD_i = \frac{1}{24}\sum_{t=1}^{24} HD_h \tag{8}$$

$$HD_h = \begin{cases} 0 & T_h < T_t \\ T_h - T_t & T_h \ge T_t \end{cases} \tag{9}$$

HD_h represents the hourly high-temperature degree days on the *i*th day. T_t indicates the temperature threshold for HTS and was set to 30 °C (Farooq et al., 2011; Liu et al., 2014). T_h is the hourly temperature. *m* in Eq. (7) is the heat stress duration.

3. Results

3.1. Model calibration and evaluation

Four crop simulation models (CERES-Wheat, Nwheat, WheatGrow, and APSIM-Wheat) were calibrated to further evaluate the model performances under extreme temperature stresses (Table S2). Generally, all models perform with satisfactory under the control temperature with NRMSE less than 20% (Table S2). However, for T2, T3, and T4 treatments of both LTS and HTS, model performances were poor with NRMSE higher than 20% (Table S3).

Table 2
Regression coefficients of observed and simulated grain protein concentration (GPC) and grain protein yield (GPY) with four wheat models under cold degree days (CDD) and heat degree days (HDD) in environment-controlled phytotron experiments.

Stress	Stage	Cultivar	Variable	Observed		CERES-Wheat		Nwheat		WheatGrow		APSIM-Wheat	
				Slope	R^2	Slope	R^2	Slope	R^2	Slope	R^2	Slope	R^2
LTS	Jointing	Yangmai16	GPC	0.0023	0.72**	0.0000	0.00	−0.0005	0.05	0.0005	0.54**	−0.0005	0.09
			GY	−0.025	0.77**	0.0000	0.00	−0.003	0.51**	0.00	0.56**	−0.002	0.12
			GPY	−0.0234	0.75**	0.0000	0.00	−0.0035	0.67**	−0.0000	0.02	−0.0021	0.11
		Xumai30	GPC	0.0027	0.49**	−0.0004	0.08	−0.0007	0.06	0.0002	0.01	−0.0008	022*
			GY	−0.023	0.82**	−0.002	0.60**	−0.004	0.60**	0.00	0.66**	−0.004	0.46**
			GPY	−0.0210	0.81**	−0.0027	0.5**	−0.0043	0.68**	−0.0003	0.03	−0.0043	0.45**
	Booting	Yangmai16	GPC	0.0032	0.72**	0.0010	0.25*	−0.0016	0.19*	0.001	0.91**	0.0000	0.01
			GY	−0.04	0.74**	0.00	0.05	−0.002	0.58**	0.00	0.91**	−0.002	0.30*
			GPY	−0.0394	0.73**	0.0003	0.01	−0.0037	0.63**	0.0003	0.53	−0.0020	0.27*
		Xumai30	GPC	0.0038	0.71**	0.0002	0.01	0.0007	0.02	0.0008	0.21*	−0.0020	0.40**
			GY	−0.046	0.75**	−0.002	0.05	−0.003	0.61**	0.00	0.78**	−0.004	0.54**
			GPY	−0.0454	0.74**	−0.0013	0.03	−0.0028	0.28**	0.0000	0.00	−0.0061	0.50**
HTS	Anthesis	Yangmai16	GPC	0.0106	0.41**	−0.0003	0.02	−0.0014	0.22	−0.009	0.82**	0.0034	0.81**
			GY	−0.011	0.82**	−0.003	0.26*	−0.005	0.88**	−0.015	0.96**	−0.008	0.78**
			GPY	−0.0032	0.05	−0.0036	0.23	−0.0065	0.85**	−0.021	0.93**	−0.0055	0.59**
		Xumai30	GPC	0.0161	0.54**	0.0000	0.00	−0.0038	0.62**	−0.010	0.79**	0.0026	0.71**
			GY	−0.024	0.91**	−0.005	0.39*	−0.002	0.69**	−0.021	0.89**	−0.012	0.86**
			GPY	−0.0173	0.72**	−0.0048	0.35*	−0.0061	0.88**	−0.026	0.89**	−0.0106	0.83**
	Grain filling	Yangmai16	GPC	0.0145	0.62**	−0.0002	0.01	0.0115	0.80**	−0.010	0.87**	0.0058	0.89**
			GY	−0.013	0.65**	0	1.00	−0.009	0.90**	−0.01	0.97**	−0.008	0.76**
			GPY	−0.0033	0.05	−0.0002	0.01	0.0005	0.02	−0.018	0.93**	−0.0037	0.27*
		Xumai30	GPC	0.0157	0.73**	−0.0002	0.01	0.0103	0.71**	−0.008	0.85**	0.0015	0.44**
			GY	−0.014	0.75**	0.00	0.32	−0.009	0.95**	−0.01	0.95**	−0.011	0.86**
			GPY	−0.0038	0.1	−0.0002	0.02	−0.0059	0.70**	−0.017	0.91**	−0.0097	0.82**
	Combined stress at anthesis and grain filling	Yangmai16	GPC	0.0121	0.59**	0.0000	0.00	0.0053	0.75**	−0.006	0.82**	0.0026	0.78**
			GY	−0.012	0.87**	−0.003	0.79**	−0.007	0.89**	−0.011	0.97**	−0.013	0.93**
			GPY	−0.0064	0.31*	−0.0030	0.62**	−0.0034	0.96**	−0.014	0.91**	−0.0118	0.87**
		Xumai30	GPC	0.0107	0.65**	−0.0003	0.05	0.0040	0.54**	−0.007	0.82**	0.0014	0.73**
			GY	−0.013	0.88**	−0.002	0.78**	−0.006	0.96**	−0.013	0.90**	−0.014	0.97**
			GPY	−0.0093	0.73**	−0.0021	0.73**	−0.0056	0.75**	−0.016	0.88**	−0.0136	0.96**

Slope and R^2 were the slopes of fitted lines and coefficients of determination between wheat yield and quality parameters and CDD or HDD.
**Indicates significance at $P < 0.01$; *indicates significance at $P < 0.05$.

3.2. Response of GPC to low temperature stress at jointing and booting stages

Low-temperature stress (LTS) at jointing and booting stages resulted in increased GPC ($P > 0.001$, Fig. 1). With per unit increase in CDD, GPC was increased by 0.2% and 0.3% at jointing and by 0.3% and 0.4% at booting for cv. Yangmai16 and cv. Xumai30, respectively. Slopes of the regression line indicated that GPC increased more at the booting stage as compared with the jointing stage under LTS (Table 2). Crop simulation models tested here failed to simulate the effect of LTS on GPC at jointing and booting stages except for WheatGrow. For example,-WheatGrow showed a positive slope with the range of 0.0–0.1% that varies with cultivar and growth stage under LTS at booting, while CERES-Wheat, Nwheat, and APSIM-Wheat showed negative slopes for GPC with increasing CDD.

3.3. Response of GPY to low temperature stress at jointing and booting stages

As shown in Fig. 2, observed wheat grain yield (GY) decreases under LTS at jointing and booting ($P > 0.001$). The per unit increase in CDD results in decrease of GY at jointing and booting by 2.5% and 4% for cv. Yangmai16 and 2.3% and 4.6% for cv. Xumai30, respectively. Slopes of the regression line showed more decrease in GY at booting than jointing (Table 2). Four simulation models fail to reproduce the impact of LTS on GY at both stages of LTS. For example, CERES-Wheat simulates no response for GY with cv. Yangmai30, but simulates the decline in GY by -0.02% at jointing and booting for cv. Xumai30. Nwheat simulates the decrease in GY at jointing and booting by 0.3% and 0.2% for cv. Yangmai16 and 0.4% and 0.3% for cv. Xumai30, respectively. WheatGrow simulates no response for GY under LTS. The range of slope for APSIM-Wheat was 0.2% to 0.4% depending on the cultivar and stage of LTS.

Grain protein yield (GPY) was more sensitive to LTS at booting than at jointing for two cultivars (Fig. 3). For example, observed GPY decreased by 2.3% and 2.1% at jointing and by 3.9% and 4.5% at booting for cv. Yangmai16 and cv. Xumai30, respectively, with per unit increase in CDD ($P > 0.001$). The four crop simulation models appeared insensitive to LTS at jointing and booting stages for GPY. The simulated decrease in slope was 0.0 to 0.3% for CERES-Wheat, 0.3% to 0.4% for Nwheat under LTS at both stages. Whereas, WheatGrow simulated different responses at different stages, with a decreased slope of 0.0% to −0.03% at jointing and showed no response at booting. APSIM-Wheat showed a decrease in the slope of 0.2% to 0.6% under LTS, and a better response was simulated for cv. Xumai30.

3.4. Response of GPC to high temperature stress at anthesis and grain filling stages

Fig. 4 shows the observed and simulated responses of GPC under HTS. Observed GPC increased significantly with high-temperature stress (HTS) at anthesis, grain filling and combined stress at anthesis and grain filling ($P < 0.001$) in both wheat cultivars. With per unit increase in HDD,the observed GPC increased by 1.1–1.6%, 1.5–1.6% and 1.1–1.2% at anthesis, grain filling and combined stress at anthesis and grain filling, depending on wheat cultivar and stage of HTS. Generally, a little higher increase in GPC occurred for cv. Xumai30 (1.1 to 1.6%) than cv. Yangmai16 (1.1 to 1.5%), especially for HTS at anthesis. Among crop simulation models, APSIM-Wheat tended to simulate some responses for GPC under HTS with a slope of 0.3% at anthesis stage, while other models either underestimated or showed no response. At grain filling stage, Nwheat simulated the best response of GPC under HTS with a slope of 1.1%, while the estimated slope for APSIM-Wheat showed a slight increase (0.4%) in GPC. In the case of HTS at combined

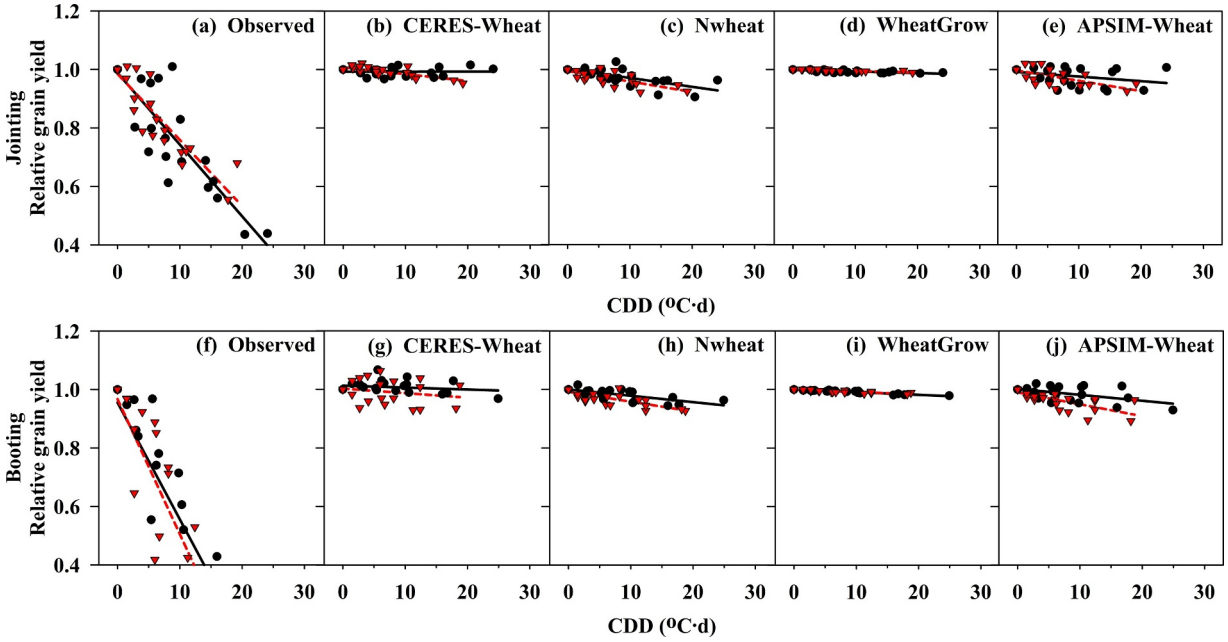

Fig. 2. Relative response of observed and simulated grain yield to low temperature stress at jointing (a–e) and booting (f–j) from four wheat models for two winter wheat cultivars in environment-controlled phytotron experiments. Observed (a and f) and simulated values with CERES-Wheat (b and g), Nwheat (c and h), WheatGrow (d and i), and APSIM-Wheat (e and j). Black circle and straight line indicate cv. Yangmai16; red triangle and dash line indicate cv. Xumai30.

stress at anthesis and grain filling, a predicted increase in GPC with Nwheat was approximately 0.4–0.5%, while the estimated slope was 0.1–0.3% with APSIM-Wheat. Nwheat showed better simulation results under thecombined stress at anthesis and grain filling than other models.

3.5. Response of GPY to high temperature stress at anthesis and grain filling stages

GY decreased more at anthesis as compared with grain filling and combined stress at anthesis and grain filling (Table 2). With every unit increase in HDD, the decrease in GY at anthesis, grain filling, and combined stress at anthesis and grain filling was 1.1%, 1.3% and 1.2% for cv. Yangmai16, and 2.4%, 1.4% and 1.3% for cv. Xumai30, respectively ($P < 0.001$). Fig. 5 shows the responses of four models (CERES-Wheat, Nwheat, WheatGrow, and APSIM-Wheat) under different levels and durations of HTS. CERES-Wheat does not simulate any responses to HTS at any stage of treatment. While Nwheat, WheatGrow, and APSIM-Wheat reproduce some effects of HTS on GY. For example, with each unit increase in HDD, Nwheat shows a decrease in GY by 0.2% to 0.9%, depending on the stages of HTS and cultivars. Among the crop models, the WheatGrow model simulates the best results for the

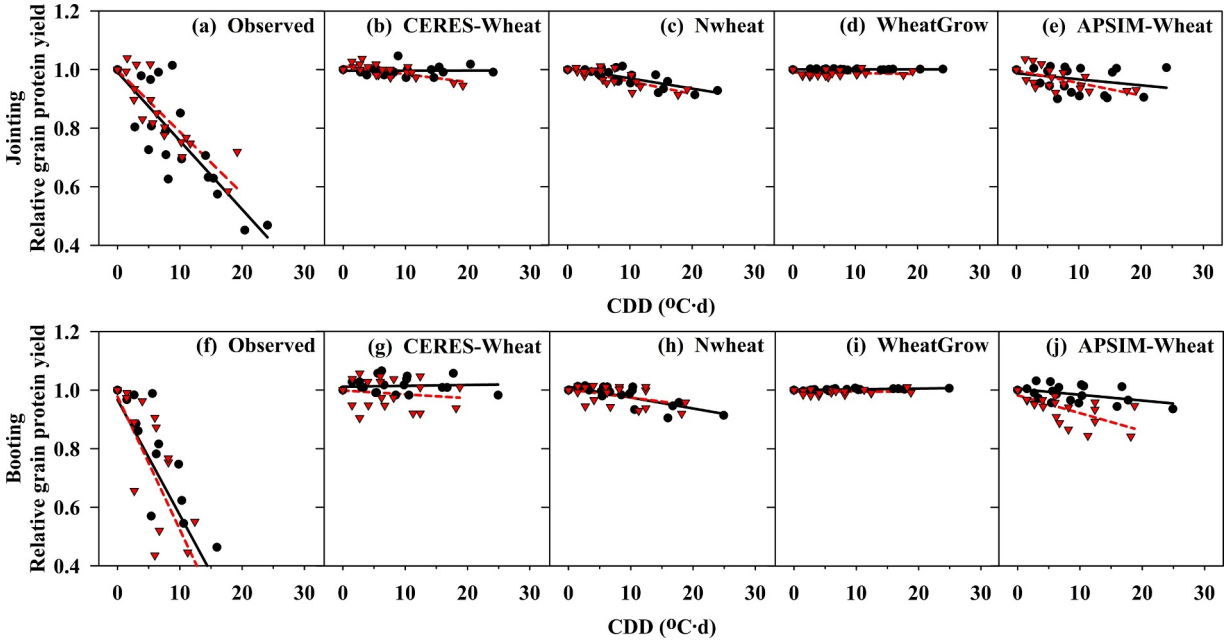

Fig. 3. Relative response of observed and simulated grain protein yield to low temperature stress at jointing (a–e) and booting (f–j) from four wheat models for two winter wheat cultivars in environment-controlled phytotron experiments. Observed (a and f) and simulated values with CERES-Wheat (b and g), Nwheat (c and h), WheatGrow (d and i), and APSIM-Wheat (e and j). Black circle and straight line indicate cv. Yangmai16; red triangle and dash line indicate cv. Xumai30.

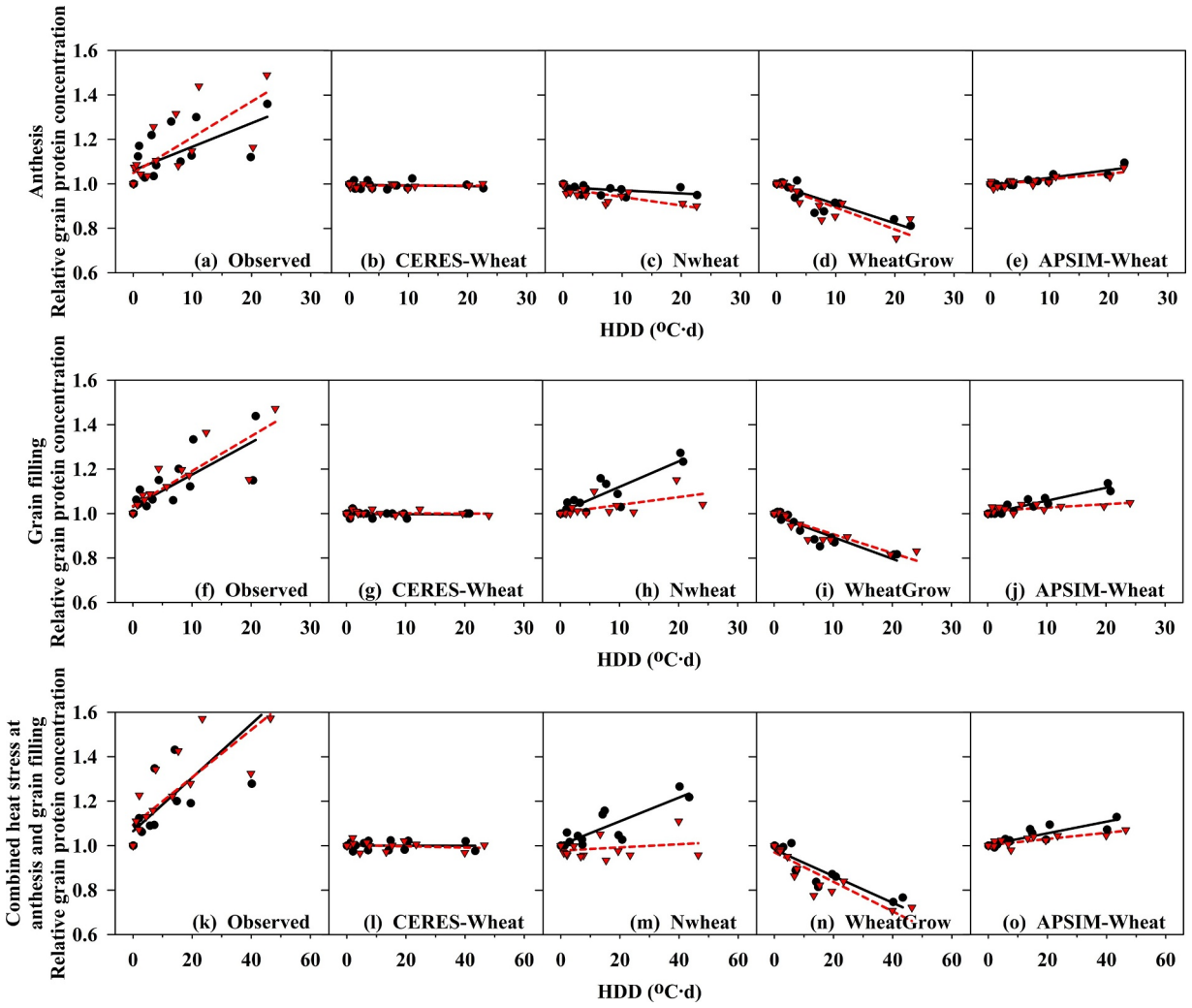

Fig. 4. Relative response of observed and simulated grain protein concentration to heat stress at anthesis (a–e), grain filling (f–j) and combined stress at anthesis and grain filling (k-o) from four wheat models for two winter wheat cultivars in environment-controlled phytotron experiments. Observed (a, f and k) and simulated values with CERES-Wheat (b, g, and l), Nwheat (c, h, and m), WheatGrow (d, i, and n), and APSIM-Wheat (e, j, and o). Black circle and straight line indicate cv. Yangmai16; red triangle and dash line indicate cv. Xumai30.

impact of HTS on GY and shows a decrease in GY from 1% to 2.1%, also varies with cultivars and the stage of treatment. APSIM-Wheat reproduces the reduction in GY under HTS by 0.8% to 1.4%, depending on the cultivars and the stage of HTS.

Observed GPY was negatively related to HTS for cv. Yangmai16 and cv. Xumai30. With every unit increment in HDD, GPY decreased approximately by 0.3% to 1.7% at anthesis, 0.3 to 0.4% at grain filling and 0.6% to 0.9% under combined stress at anthesis and grain filling for cv. Yangmai16 and cv. Xumai30, respectively. The decrease in GPY under heat stress at anthesis, grain filling and combined stress at anthesis and grain filling were more obvious than the increase in GPC. Among stages, GPY appeared to be more sensitive to HTS at anthesis, as more decrease of GPY was observed at anthesis as compared with grain filling. The observed and simulated responses of GPY to four different temperature levels (Tmin/Tmax: 17/27, 21/31, 25/35, 29/39 °C) are presented in Fig. 6. Nwheat, WheatGrow, and APSIM-Wheat showed the decreasing trend of GPY with increasing HDD, while no response was shown in CERES-Wheat. In Nwheat, WheatGrow, and APSIM-Wheat, the simulated GPY decreased 0.3% to 0.7%, 1.4% to 2.6%, and 0.4% to 1.4% with every unit increase in HDD, depending on cultivars and stages at which stress occurred. WheatGrow overestimated the decrease in GPY under HTS at all stages. While APSIM-Wheat underestimated responses of GPY at anthesis and overestimated the reductions of GPY at grain filling and combined stress at anthesis and grain filling (Table 2). Nwheat produced satisfactory results under HTS at anthesis and grain filling stages but underestimated GPY was observed under combined stress at anthesis and grain filling.

4. Discussion

Increased magnitude and frequency of extreme climatic events are major threats to global food security (Nuttall et al., 2017; Rötter et al., 2018; Trnka et al., 2014). The upper and lower temperature thresholds for wheat are >30 °C for the reproductive stages (Farooq et al., 2011; Liu et al., 2014) and <2 °C (Crimp et al., 2016; Ji et al., 2017) for the vegetative stages. Unusual responses in both GY and grain quality are expected under climatic extremes.

Low temperature appears to be one of the main environmental extremes that reduce wheat yield (Ji et al., 2017). The effects of LTS on wheat are determined by duration, level and growth stage at which stress occurs (Craufurd et al., 2013; Porter and Gawith, 1999). The significant differences of GY and GPC between cultivars, temperatures levels and durations under LTS in this study were observed by Ji et al. (2017) and Liu et al. (2019). The present study showed the increasing GPC, and decreasing GY and GPY with decreasing temperature levels under LTS. GPC, GY, and GPY seem to be more sensitive

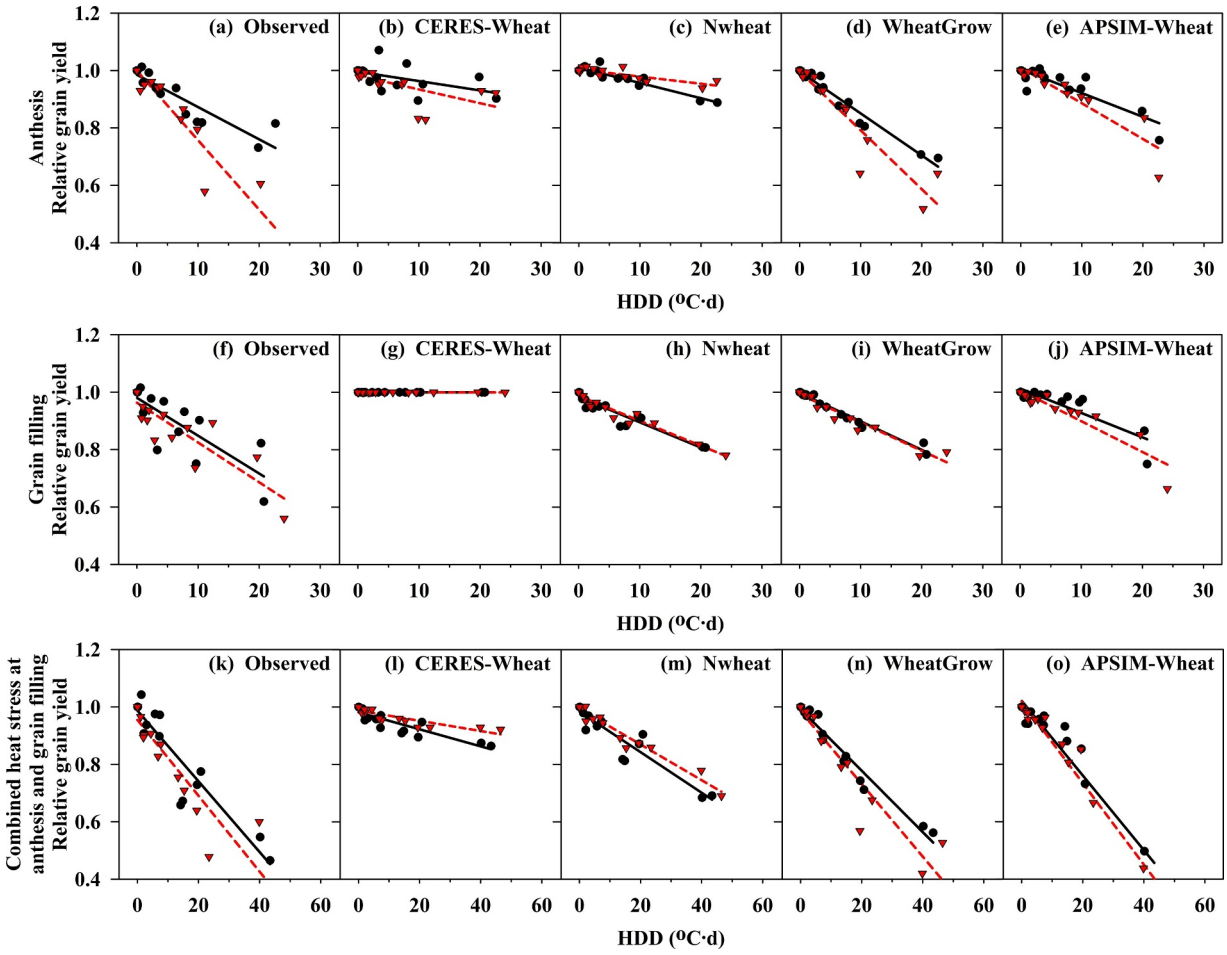

Fig. 5. Relative response of observed and simulated grain yield to heat stress at anthesis (a–e), grain filling (f–j) and combined stress at anthesis and grain filling (k-o) from four wheat models for two winter wheat cultivars in environment-controlled phytotron experiments. Observed (a, f, and k) and simulated values with CERES-Wheat (b, g, and l), Nwheat (c, h, and m), WheatGrow (d, i, and n), and APSIM-Wheat (e, j, and o). Black circle and straight line indicate cv. Yangmai16; and red triangle and dash line indicate cv. Xumai30.

to LTS at booting than at jointing stage. As young spikes at booting stage had less resistance to LTS, and further prolonged duration of LTS (>0 °C) during pollen growth and development can result in aborted florets, pollen sterility and incomplete ear structure (Frederiks et al., 2015; Ji et al., 2017). However, LTS at jointing and booting can cause a decrease in the maximum quantum efficiency of PSII and gas exchange rates, which reduce the photosynthetic carbon assimilation in leaves and further cause reduced biomass accumulation (Li et al., 2015). Thus, the reduced biomass before anthesis may have a negative effect on the differentiation of spikelet. As suggested by Ji et al. (2017), among spike number per plant, number of grain per spike, and thousand grain weight, spike number per plant and number of grains per spike can be affected by LTS at jointing and booting stages. In addition, LTS at different critical stages may also have different effects on plant N dynamics, which also could affect final GPC and GPY. However, less study is available currently and should be conducted in the future.

The observed increasing GPC and decreasing GY and GPY under HTS (Fig. 4-6) was consistent with previous results from both environmentally controlled phytotron and field environments (Asseng and Milroy, 2006). According to Asseng and Milroy (2006), the occurrence of HTS at anthesis and grain filling stages resulted in higher GPC than in controlled treatments, due to the different temperature responses of dry matter and N accumulation rate in the grain. Grain starch deposition decreases or ceases if temperature exceeds beyond the threshold level (>30 °C) due to reduced activity of starch synthesis enzymes (Nuttall et al., 2017). On the contrary, most of grain protein deposition is unaffected by HTS (Nuttall et al., 2017), and an increase in GPC may result (Asseng and Milroy, 2006; Liu et al., 2016a; Nuttall et al., 2017; Panozzo and Eagles, 2000). This indicated that the increased GPC was due to that the reduction in grain mass was higher than that of protein under HTS (Farooq et al., 2011; Liu et al., 2016a; Nuttall et al., 2017). Because GPY was determined by GY and GPC, a decrease in GPY was a result of a higher decrease in GY than increase in GPC, as shown in our data here (Table 2).

Crop simulation models appear to be a useful tool to measure the climate impact on crop productivity (Asseng et al., 2015), but unfortunately, most of the contemporary crop models were found to be insensitive under extreme temperature events (LTS and HTS) to some extent (Liu et al., 2016a; Rötter et al., 2011). Most of recent model improvements conducted in wheat mostly focused on HTS impacts on wheat growth and yield (Maiorano et al., 2017; Wang et al., 2017; Webber et al., 2018), while less on the impacts of LTS, and also less on the impacts of extreme temperature on grain quality. Generally, current wheat models could correctly simulate GPY and GPC without any stress treatment, but the model performance is still needed to improve under extreme temperature conditions. Our results of model comparison for simulating GPC and GPY corroborate with previous investigations (Barlow et al., 2015; Liu et al., 2016a; Lobell et al., 2012; Moriondo et al., 2011), as these models usually were developed and evaluated under normal climate conditions (Angulo et al., 2013). For example, Nwheat and APSIM-Wheat showed a response to HTS for GPC, but under LTS all models failed to simulate GPC. For GY, all models fail

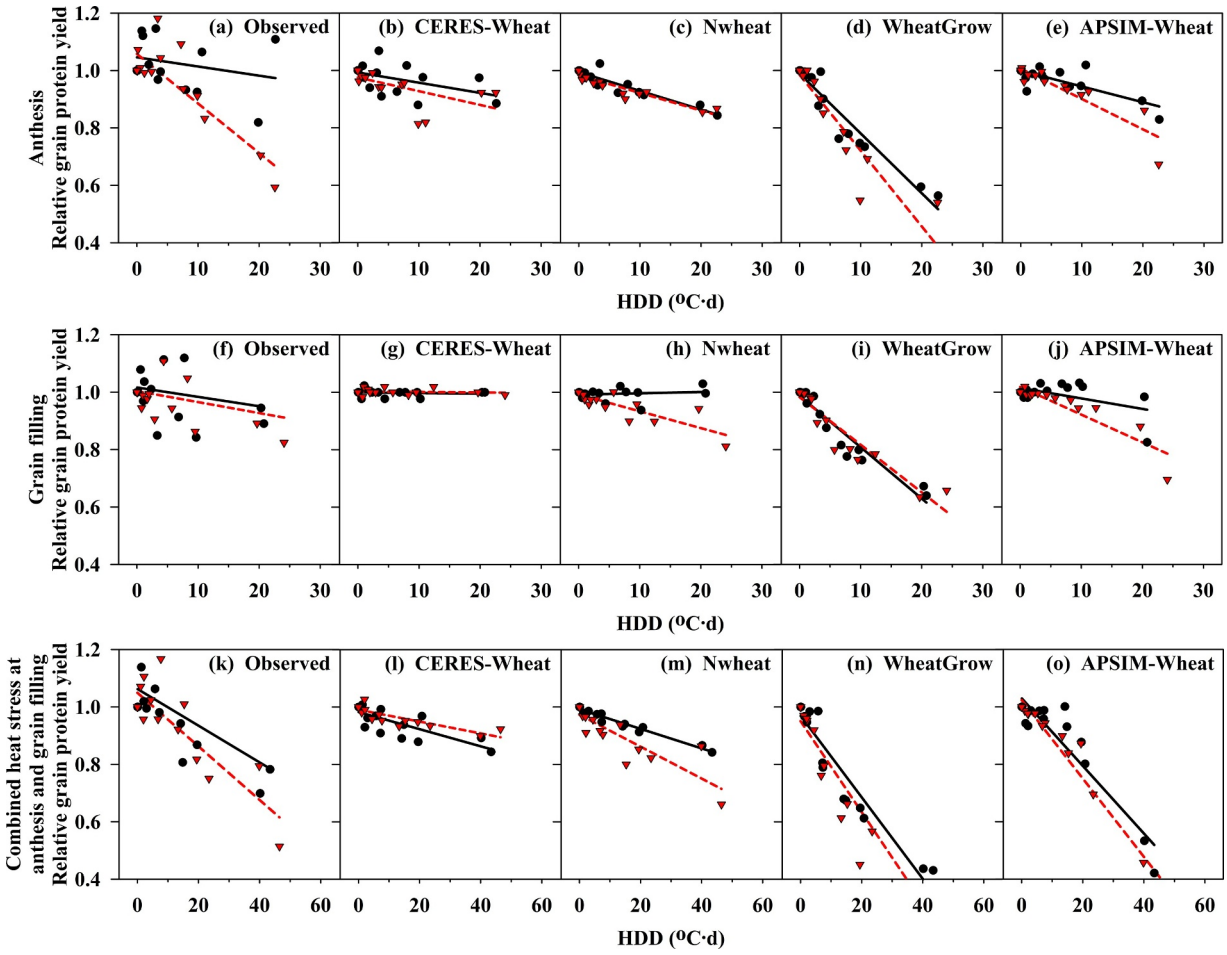

Fig. 6. Relative response of observed and simulated grain protein yield to heat stress at anthesis (a–e), grain filling (f–j) and combined stress at anthesis and grain filling (k-o) from four wheat models for two winter wheat cultivars in environment-controlled phytotron experiments. Observed (a, f, and k) and simulated values with CERES-Wheat (b, g, and l), Nwheat (c, h, and m), WheatGrow (d, i, and n), and APSIM-Wheat (e, j, and o). Black circle and straight line indicate cv. Yangmai16; red triangle and dash line indicate cv. Xumai30.

to simulate the impact of LTS. However, except for CERES-Wheat, other models like Nwheat, WheatGrow, and APSIM-Wheat, simulates some impacts of HTS on GY. In the case of GPY, Nwheat, WheatGrow, and APSIM-Wheat produced a reduction under HTS, but CERES-Wheat showed no response. Similarly, Nwheat and APSIM-Wheat predicted some reductions in GPY under LTS, but WheatGrow and CERES-Wheat failed to produce a reduction in GPY. Nwheat and APSIM-Wheat show a limited response in increasing GPC under HTS, because they simulate the impact of HTS on grain yield but did not consider the effects of HTS on grain N accumulation. According to Asseng et al. (2011), Nwheat accelerates leaf senescence with temperature above 34 °C which decreases the duration of starch accumulation in grains and results in the increased GPC and decreased GPY. WheatGrow reproduces the impact of HTS on GY but did not simulate the increase in GPC. This is due to the fact that WheatGrow uses separate modules for GY and grain N accumulation. Several improvements have been made for incorporating HTS impacts on wheat growth, biomass accumulation, and yield formation in WheatGrow, but in the original simulations of grain N, the rate of grain N accumulation will be reduced after 24.2 °C (Pan et al., 2006), which may result in a decrease in grain N under HTS. Therefore, the improvement of WheatGrow is needed in the near future for N dynamics and GPC. APSIM-Wheat simulates less increase in GPC and more reduction in GPY under HTS treatments with respect to Nwheat. This may be due to an under estimation of GPC under HTS in APSIM-Wheat (Liu et al., 2016a).

Crop simulation models provide a good insight into physiological processes that occur within plants. Four extensively used wheat simulation models were evaluated against detailed observed responses of grain quality indices from a range of extreme temperature stress treatments. The response of the four crop models to extreme temperature stress varied significantly in the simulations of GPC and GPY. Even some models in this study gave some responses to some extent under extreme temperatures, improvement is still needed of temperature response functions for extreme temperature stress to adapt to future climate scenarios (Barlow et al., 2015; Nuttall et al., 2017). Therefore, Liu et al. (2016b) and Liu et al. (2017) have improved the simulations for phenology and biomass accumulation under HTS by incorporating high-temperature responses in WheatGrow model, but the model improvements of grain quality are still needed under extreme temperature stresses.

Hence, Orlando et al. (2017) emphasized the need to review the current algorithms for GPC modeling, especially under extreme temperature stresses. Several simulation models have included the effect of N on wheat growth, development and yield formation (Jamieson and Semenov, 2000), and plant N demand and soil N availability normally can be used to simulate GPC in some models, such as APSIM-Nwheat (Asseng and Milroy, 2006), CERES-wheat (Ritchie, 1985), AFRCWHEAT2 (Porter, 1993) and SWHEAT and SIRIUS (Van Keulen and Seligman, 1987). However, few of them simulated the effect of temperature on N uptake from soil and N translocation within plant, especially N translocation to grain. In addition, less impact of extreme temperature on grain N accumulation were considered currently (Table

S1). Incorporating the extreme temperature effects along with the genotypic parameters into the simulations of plant N dynamics for correctly predicting grain N concentration (e.g. N uptake and translocation to grains) will be critical for improving model performance on GPC modeling under extreme temperatures. Due to the differences in extreme temperature sensitivity at different growth stages, stage-dependent temperature threshold are required for N uptake and translocation.

Generally, simulations of GPC and GPY are dependent on the simulations of grain yield in most of the crop models. Biased estimation of GY will also lead to biased simulations of grain quality. Therefore, improving simulations of both grain N accumulation and GY under extreme temperature stresses will be necessary for current wheat models, especially under LTS, because current models were unable to reproduce the effects of LTS on GY and yield components (Fig. 2 and S4-S5). For HTS, except CERES-Wheat, other models e.g., Nwheat, WheatGrow and APSIM-Wheat, show some responses for GY under HTS. However, these models show large variations in simulating GY and yield components under HTS (Fig. 5 and S6-S7). For example, Nwheat shows the decrease in simulated grain number at anthesis, grain filling, and combined stress at anthesis and grain filling, but still underestimates the impact of HTS on grain size and grain yield. WheatGrow simulates the reduction in grain number at anthesis and combined stress at anthesis and grain filling, but shows no response at the grain filling. APSIM-Wheat was unable to produce satisfactory results for the impact of HTS on grain number for all stages, because grain number in APSIM-Wheat is determined by stem weight at anthesis (Zadoks 65),and was less affected by HTS. However, in the case of grain size, APSIM-Wheat simulates well the reduction in grain size for all stages under HTS. Here, some models (e.g. Nwheat, WheatGrow, and APSIM-Wheat) tended to reproduce some impacts of HTS on GPC and GPY, but these models didn't incorporate any specific algorithms for extreme temperatures on grain N to improve the simulations of grain quality. This could be due to the improvements in the impacts of HTS on the formation of GY in these models. Therefore, better simulations of GY could result in better simulations of GPC and GY in existing crop models.

According to Siebert et al. (2014), the mean decline of wheat yield due to HTS during the anthesis was 0.7% when the temperatures were calculated at a height of 2 m, however, the yield decline was 22% for temperatures measured at ground. These uncertainties among air and canopy temperature results in biased simulations of GPC and GPY. The differences between canopy temperature and air temperature are mainly associated with plant water status and environmental factors (Siebert et al., 2014; Webber et al., 2017). Generally, water deficit can increase canopy temperature and may exacerbate HTS impacts on crops, due to reduced transpiration rates (Siebert et al., 2014). In contrast, crop canopy temperature are almost lower than air temperature under irrigated conditions, especially when vapor pressure deficit (VPD) is high (Kimball et al., 2012; Siebert et al., 2017), and high temperature responses developed under current climate conditions using air temperature are not expected to be the same under higher temperature with climate change (Siebert et al., 2017). As higher temperature usually has exponentially higher VPD and therefore evaporative cooling effect. In our experiments, full irrigation was applied to avoid any water logging or drought condition, and air temperature was used to quantify the impacts of LTS and HTS for this study. Thus, the estimated responses of grain yield and quality parameters to extreme temperature stress may need to be reconsidered for applications in other environment conditions. Several recent studies concluded that irrigation may help to reduce the adverse effects of extreme temperatures on crop yield (Carter et al., 2016; Troy et al., 2015), and emphasized the use of canopy temperature for estimating temperature impacts across different environments and production conditions to improve the predictability of crop models (Webber et al., 2018; Yoshimoto et al., 2011). Therefore, future studies should consider the interactions between extreme temperature stress (both LTS and HTS) and water conditions in model improvements, especially for model applications under warming scenarios, to reduce the uncertainties in assessing extreme temperature impacts on crop yield and quality due to without considering the dynamics of canopy temperature (Siebert et al., 2017).

Further, high tolerance to extreme temperature was suggested as one of crop adaptation strategies to future climate change (Gouache et al., 2012; Zheng et al., 2012), thus model improvements will also need to account for cultivar-specific tolerance to extreme temperatures, which has been demonstrated in the two tested cultivars of this study. As an example, Stratonovitch and Semenov (2015) also have taken genotypic differences in an account under HTS when adapting wheat production to future climate impacts in Europe. In addition, most of the previous adaptation studies only focused on the strategies for GY improvement. However, quantification of the effects of adaptation strategies on grain quality will be critical in the near future, because some of the adaptation strategies to future climate scenarios may have adverse effects on grain protein formation (Asseng et al., 2019).

Declaration of Competing Interest

The authors declare no competing interests.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2019YFA0607404), the National Science Fund for Distinguished Young Scholars (31725020), the Fundamental Research Funds for the Central Universities (KJQN201902), the National Natural Science Foundation of China (31872848, 31801260, 41961124008), the Natural Science Foundation of Jiangsu province (BK20180523), the Young Elite Scientists Sponsorship Program by CAST (2017QNRC001), and the China Scholarship Council.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.agrformet.2020.107995.

References

Alderman, P.D., Quilligan, E., Asseng, S., Ewert, F. and Reynolds, M.P., 2013. Proceedings of the Workshop Modeling Wheat Response to High Temperature held in El Batán, Mexico, 19-21 June 2013.

Angulo, C., 2013. . Implication of crop model calibration strategies for assessing regional impacts of climate change in Europe. Agr. Forest Meteorol. 170, 32–46.

Asseng, S., 2002. . Simulation of grain protein content with APSIM-Nwheat. Eur. J. Agron. 16 (1), 25–42.

Asseng, S., 2015. Rising temperatures reduce global wheat production. Nat. Clim. Chang. 5 (2), 143.

Asseng, S., Foster, I.A.N., Turner, N.C., 2011. The impact of temperature variability on wheat yields. Glob. Chang. Biol. 17 (2), 997–1012.

Asseng, S., 2019. . Climate change impact and adaptation for wheat protein. Global Change Biol. 25 (1), 155–173.

Asseng, S., Milroy, S.P., 2006. Simulation of environmental and genetic effects on grain protein concentration in wheat. Eur. J. Agron. 25 (2), 119–128.

Augspurger, C.K., 2013. Reconstructing patterns of temperature, phenology, and frost damage over 124 years: spring damage risk is increasing. Ecology 94 (1), 41–50.

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. Field Crop Res. 171, 109–119.

Carter, E.K., Melkonian, J., Riha, S.J., Shaw, S.B., 2016. Separating heat stress from moisture stress: analyzing yield response to high temperature in irrigated maize. Environ. Res. Lett. 11 (9), 094012.

Chenu, K., 2017. . Contribution of Crop Models to Adaptation in Wheat. Trends Plant Sci. 22 (6), 472–490.

Craufurd, P.Q., Vadez, V., Jagadish, S.V.K., Prasad, P.V.V., Zaman-Allah, M., 2013. Crop science experiments designed to inform crop modeling. Agric. For. Meteorol. 170, 8–18.

Crimp, S., 2015. . Bayesian space–time model to analyse frost risk for agriculture in Southeast Australia. Int. J. Climatol. 35 (8), 2092–2108.
Crimp, S.J., 2016. . Recent changes in southern Australian frost occurrence: implications for wheat production risk. Crop Pasture Sci. 67 (8), 801–811.
Dias, A.S., Lidon, F.C., 2009. Evaluation of grain filling rate and duration in bread and durum wheat, under heat stress after anthesis. J. Agron. Crop Sci. 195 (2), 137–147.
DuPont, F.M., Altenbach, S.B., 2003. Molecular and biochemical impacts of environmental factors on wheat grain development and protein synthesis. J. Cereal Sci. 38 (2), 133–146.
Farooq, M., Bramley, H., Palta, J.A., Siddique, K.H.M., 2011. Heat stress in wheat during reproductive and grain-filling phases. CRC Crit. Rev. Plant Sci. 30 (6), 491–507.
Frederiks, T.M., Christopher, J.T. and Borrell, A.K., 2008. Low temperature adaption of wheat post head-emergence in northern Australia.
Frederiks, T.M., Christopher, J.T., Sutherland, M.W., Borrell, A.K., 2015. Post-head-emergence frost in wheat and barley: defining the problem, assessing the damage, and identifying resistance. J. Exp. Bot. 66 (12), 3487–3498.
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 (4), 435–441.
Gouache, D., 2012. . Evaluating agronomic adaptation options to increasing heat stress under climate change during wheat grain filling in France. Eur. J. Agron. 39, 62–70.
Gu, L., 2008. . The 2007 eastern US spring freeze: increased cold damage in a warming world? Bioscience 58 (3), 253–262.
Jamieson, P.D., Semenov, M.A., 2000. Modelling nitrogen uptake and redistribution in wheat. Field Crop Res. 68 (1), 21–29.
Ji, H., 2017. . Effects of jointing and booting low temperature stresses on grain yield and yield components in wheat. Agric. For. Meteorol. 243, 33–42.
Kassie, B.T., Asseng, S., Porter, C.H., Royce, F.S., 2016. Performance of DSSAT-Nwheat across a wide range of current and future growing conditions. Eur. J. Agron. 81, 27–36.
Kawasaki, K., Uchida, S., 2016. Quality Matters more than quantity: asymmetric temperature effects on crop yield and quality grade. Am. J. Agric. Econ. 98 (4), 1195–1209.
Keating, B.A., 2003. . An overview of APSIM, a model designed for farming systems simulation. Eur. J. Agron. 18 (3–4), 267–288.
Kimball, B.A., White, J.W., Wall, G.W., Ottman, M.J., 2012. Infrared-warmed and unwarmed wheat vegetation indices coalesce using canopy-temperature–based growing degree days. Agron. J. 104 (1), 114–118.
Kodra, E., Steinhaeuser, K., Ganguly, A.R., 2011. Persisting cold extremes under 21st-century warming scenarios. Geophys. Res. Lett. 38 (8).
Li, X., 2015. . Winter wheat photosynthesis and grain yield responses to spring freeze. Agron. J. 107 (3), 1002–1010.
Liu, B., 2016a. . Testing the responses of four wheat crop models to heat stress at anthesis and grain filling. Global Change Biol. 22 (5), 1890–1903.
Liu, B., 2017. . Modelling the effects of post-heading heat stress on biomass growth of winter wheat. Agr. Forest Meteorol. 247, 476–490.
Liu, B., 2016b. . Modelling the effects of heat stress on post-heading durations in wheat: a comparison of temperature response routines. Agric. For. Meteorol. 222, 45–58.
Liu, B., 2014. . Post-heading heat stress and yield impact in winter wheat of China. Glob. Chang. Biol. 20 (2), 372–381.
Liu, L., 2019. . Response of wheat grain quality to low temperature during jointing and booting stages-On the importance of considering canopy temperature. Agric. For. Meteorol. 278, 107658.
Lobell, D.B., Sibley, A., Ortiz-Monasterio, J.I., 2012. Extreme heat effects on wheat senescence in India. Nat. Clim. Chang. 2 (3), 186.
Lv, Z., Liu, X., Cao, W., Zhu, Y., 2013. Climate change impacts on regional winter wheat production in main wheat production regions of China. Agr. Forest Meteorol. 171, 234–248.
Maiorano, A., 2017. . Crop model improvement reduces the uncertainty of the response to temperature of multi-model ensembles. Field Crops Res. 202, 5–20.
Martre, P., Porter, J.R., Jamieson, P.D., Triboï, E., 2003. Modeling grain nitrogen accumulation and protein composition to understand the sink/source regulations of nitrogen remobilization for wheat. Plant Physiol. 133 (4), 1959–1967.
Moriondo, M., Giannakopoulos, C., Bindi, M., 2011. Climate change impact assessment: the role of climate extremes in crop yield simulation. Clim. Change 104 (3–4), 679–701.
Nuttall, J.G., 2017. . Models of grain quality in wheat-A review. Field Crop Res. 202, 136–145.
O'Leary, G.J., 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 (7), 2670–2686.
Orlando, F., 2017. Modelling durum wheat (Triticum turgidum L. var. durum) grain protein concentration. J. Agric. Sci. 155 (6), 930–938.
Palosuo, T., 2011. . Simulation of winter wheat yield and its variability in different climates of Europe: a comparison of eight crop growth models. Eur. J. Agron. 35 (3), 103–114.
Pan, J., 2006. . Modeling plant nitrogen uptake and grain nitrogen accumulation in wheat. Field Crop Res. 97 (2–3), 322–336.
Panozzo, J.F., Eagles, H.A., 2000. Cultivar and environmental effects on quality characters in wheat. II. Protein. Aust. J. Agricult. Res. 51 (5), 629–636.
Peings, Y., Cattiaux, J., Douville, H., 2013. Evaluation and response of winter cold spells over Western Europe in CMIP5 models. Clim. Dyn. 41 (11–12), 3025–3037.
Porter, J.R., 1993. AFRCWHEAT2: a model of the growth and development of wheat incorporating responses to water and nitrogen. Eur. J. Agron. 2 (2), 69–82.
Porter, J.R., Gawith, M., 1999. Temperatures and the growth and development of wheat: a review. Eur. J. Agron. 10 (1), 23–36.
Potithep, S., Yasuoka, Y., 2011. Application of the 3-PG model for gross primary productivity estimation in deciduous broadleaf forests: a study area in Japan. Forests 2 (2), 590–609.
Pradhan, G.P., Prasad, P.V.V., Fritz, A.K., Kirkham, M.B., Gill, B.S., 2012. Effects of drought and high temperature stress on synthetic hexaploid wheat. Funct. Plant Biol. 39 (3), 190–198.
Ritchie, J.T., 1985. Description and performance of CERES wheat: a user-oriented wheat yield model. ARS wheat yield project 159–175.
Rötter, R.P., 2018. . Linking modelling and experimentation to better capture crop impacts of agroclimatic extremes-A review. Field Crops Res. 221, 142–156.
Rötter, R.P., Carter, T.R., Olesen, J.E., Porter, J.R., 2011. Crop-climate models need an overhaul. Nat. Clim. Chang. 1 (4), 175.
Rötter, R.P., 2012. . Simulation of spring barley yield in different climatic zones of Northern and Central Europe: a comparison of nine crop models. Field Crop Res. 133, 23–36.
Sánchez, B., Rasmussen, A., Porter, J.R., 2014. Temperatures and the growth and development of maize and rice: a review. Global Change Biol. 20 (2), 408–417.
Siebert, S., Ewert, F., Rezaei, E.E., Kage, H., Graß, R., 2014. Impact of heat stress on crop yield-on the importance of considering canopy temperature. Environ. Res. Lett. 9 (4), 044012.
Siebert, S., Webber, H., Zhao, G., Ewert, F., 2017. Heat stress is overestimated in climate impact studies for irrigated agriculture. Environ. Res. Lett. 12 (5), 054023.
Spiertz, J.H.J., 2006. . Heat stress in wheat (Triticum aestivum L.): effects on grain growth and quality traits. Eur. J. Agron. 25 (2), 89–95.
Stratonovitch, P., Semenov, M.A., 2015. Heat tolerance around flowering in wheat identified as a key trait for increased yield potential in Europe under climate change. J. Exp. Bot. 66 (12), 3599–3609.
Teixeira, E.I., Fischer, G., Van Velthuizen, H., Walter, C., Ewert, F., 2013. Global hot-spots of heat stress on agricultural crops due to climate change. Agric. For. Meteorol. 170, 206–215.
Timsina, J., Humphreys, E., 2006. Performance of CERES-Rice and CERES-Wheat models in rice–wheat systems: a review. Agric. Syst. 90 (1–3), 5–31.
Trnka, M., 2014. . Adverse weather conditions for European wheat production will become more frequent with climate change. Nat. Clim. Change 4 (7), 637.
Troy, T.J., Kipgen, C., Pal, I., 2015. The impact of climate extremes and irrigation on US crop yields. Environ. Res. Lett. 10 (5), 054013.
Van Keulen, H., Seligman, N.G., 1987. Simulation of water use, nitrogen nutrition and growth ofa spring wheat crop. Simulation Monographs. Pudoc, Wageningen, The Netherlands.
Wang, E., 2017. . The uncertainty of crop yield projections is reduced by improved temperature response functions. Nat. Plants 3 (8), 17102.
Wardlaw, I.F., 2002. Interaction between drought and chronic high temperature during kernel filling in wheat in a controlled environment. Ann. Bot. 90 (4), 469–476.
Webber, H., 2017. . Canopy temperature for simulation of heat stress in irrigated wheat in a semi-arid environment: a multi-model comparison. Field Crop Res. 202, 21–35.
Webber, H., 2018. . Physical robustness of canopy temperature models for crop heat stress simulation across environments and production conditions. Field Crops Res. 216, 75–88.
Whaley, J.M., Kirby, E.J.M., Spink, J.H., Foulkes, M.J., Sparkes, D.L., 2004. Frost damage to winter wheat in the UK: the effect of plant population density. Eur. J. Agron. 21 (1), 105–115.
Williams, S., 1984. AOAC official methods of analysis. Association of Official Analytical Chemists, 14th Edition, Arlington, VA: 8–34.
Yin, X., 2017. . Multi-model uncertainty analysis in predicting grain N for crop rotations in Europe. Eur. J. Agron. 84, 152–165.
Yoshimoto, M., 2011. . Integrated micrometeorology model for panicle and canopy temperature (IM2PACT) for rice heat stress studies under climate change. J. Agricult. Meteorol. 67 (4), 233–247.
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 (12), 3611–3623.
Zheng, B., Chenu, K., Doherty, A., Chapman, S., 2014. The APSIM-wheat Module (7.5 R3008). Agricultural Production Systems Simulator (APSIM) Initiative, Toowoomba, Australian.
Zheng, B., Chenu, K., Fernanda Dreccer, M, Chapman, S.C., 2012. Breeding for the future: what are the potential impacts of future frost and heat events on sowing and flowering time requirements for A ustralian bread wheat (Triticum aestivum) varieties? Glob. Chang. Biol. 18 (9), 2899–2914.