UNIVERSITY OF LEEDS

This is a repository copy of *The uncertainty of crop yield projections is reduced by improved temperature response functions*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/119376/

Version: Accepted Version

Article:

Wang, E, Martre, P, Zhao, Z et al. (53 more authors) (2017) The uncertainty of crop yield projections is reduced by improved temperature response functions. Nature Plants, 3. 17102. ISSN 2055-026X

https://doi.org/10.1038/nplants.2017.102

(c) 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved. This is an author produced version of a paper published in Nature Plants. Uploaded in accordance with the publisher's self-archiving policy.

Reuse
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher’s website.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

The uncertainty of crop yield projections is reduced by improved temperature response functions

Enli Wang[1]*, Pierre Martre[2]*, Zhigan Zhao[3,1], Frank Ewert[4]†, Andrea Maiorano[2]‡, Reimund P. Rötter[5,6]¶, Bruce A. Kimball[7], Michael J. Ottman[8], Gerard W. Wall[7], Jeffrey W. White[7], Matthew P. Reynolds[9], Philip D. Alderman[9]||§, Pramod K. Aggarwal[10], Jakarat Anothai[11], Bruno Basso[12], Christian Biernath[13], Davide Cammarano[14]¥, Andrew J. Challinor[15,16], Giacomo De Sanctis[17], Jordi Doltra[18], Elias Fereres[19,20], Margarita Garcia-Vila[19,20], Sebastian Gayler[21], Gerrit Hoogenboom[11] ϕ, Leslie A. Hunt[22], Roberto C. Izaurralde[23,24], Mohamed Jabloun[25], Curtis D. Jones[23], Kurt C. Kersebaum[26], Ann-Kristin Koehler[15], Leilei Liu[27], Christoph Müller[28], Soora Naresh Kumar[29], Claas Nendel[26], Garry O'Leary[30], Jørgen E. Olesen[25], Taru Palosuo[31], Eckart Priesack[13], Ehsan Eyshi Rezaei[4], Dominique Ripoche[32], Alex C. Ruane[33], Mikhail A. Semenov[34], Iurii Shcherbak[12]€, Claudio Stöckle[35], Pierre Stratonovitch[34], Thilo Streck[21], Iwan Supit[36], Fulu Tao[31,37], Peter Thorburn[38], Katharina Waha[28]¢, Daniel Wallach[39], Zhimin Wang[3], Joost Wolf[36], Yan Zhu[27] and Senthold Asseng[14]

[1]CSIRO Agriculture & Food, Black Mountain ACT 2601, Australia.
[2] UMR LEPSE, INRA, Montpellier SupAgro, 2 place Viala, 34 060, Montpellier, France.
[3]College of Agronomy and Biotechnology, China Agricultural University, Beijing, 100193, China.
[4]Institute of Crop Science and Resource Conservation (INRES), University of Bonn, 53115, Germany.
[5]University of Goettingen, Tropical Plant Production and Agricultural Systems Modelling (TROPAGS), Department of Crop Sciences, Georg-August-Universität Göttingen, 37077 Göttingen, Germany
[6]University of Goettingen, Centre of Biodiversity and Sustainable Land Use (CBL), Büsgenweg 1, 37077 Göttingen, Germany.
[7]USDA, Agricultural Research Service, U.S. Arid-Land Agricultural Research Center, Maricopa, AZ 85138, USA.
[8]The School of Plant Sciences, University of Arizona, Tucson, AZ 85721, USA.
[9]Global Wheat Program, International Maize and Wheat Improvement Center (CIMMYT) Apdo, 06600 Mexico, D.F., Mexico.
[10]CGIAR Research Program on Climate Change, Agriculture and Food Security, Borlaug Institute for South Asia, International Maize and Wheat Improvement Center (CIMMYT), New Delhi 110012, India.
[11]AgWeatherNet Program, Washington State University, Prosser, WA 99350-8694, USA.
[12]Department of Earth and Environmental Sciences and W.K. Kellogg Biological Station, Michigan State University East Lansing, MI 48823, USA.
[13]Institute of Biochemical Plant Pathology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, 85764, Germany
[14]Agricultural and Biological Engineering Department, University of Florida, Gainesville, FL 32611, USA.
[15]Institute for Climate and Atmospheric Science, School of Earth and Environment, University of Leeds, Leeds LS29JT, UK.
[16]CGIAR-ESSP Program on Climate Change, Agriculture and Food Security, International Centre for Tropical Agriculture (CIAT), A.A. 6713, Cali, Colombia.
[17]European Food Safety Authority (EFSA), GMO Unit, Parma, Italy
[18]Cantabrian Agricultural Research and Training Centre (CIFA), 39600 Muriedas, Spain.
[19]Dep. Agronomia, University of Cordoba, Apartado 3048, 14080 Cordoba, Spain.
[20]IAS-CSIC, Cordoba 14080, Spain.
[21]Institute of Soil Science and Land Evaluation, University of Hohenheim, 70599 Stuttgart.
[22]Department of Plant Agriculture, University of Guelph, Guelph, ON N1G 2W1, Canada.
[23]Dept. of Geographical Sciences, Univ. of Maryland, College Park, MD 20742, USA.

[24]Texas A&M AgriLife Research and Extension Center, Texas A&M University, Temple, TX 76502, USA.
[25]Department of Agroecology, Aarhus University, 8830 Tjele, Denmark.
[26]Institute of Landscape Systems Analysis, Leibniz Centre for Agricultural Landscape Research, 15374 Müncheberg, Germany.
[27]National Engineering and Technology Center for Information Agriculture, Key Laboratory for Crop System Analysis and Decision Making, Ministry of Agriculture, Jiangsu Key Laboratory for Information Agriculture, Jiangsu Collaborative Innovation Center for Modern Crop Production, Nanjing Agricultural University, Nanjing, Jiangsu 210095, China.
[28]Potsdam Institute for Climate Impact Research, 14473 Potsdam, Germany.
[29]Centre for Environment Science and Climate Resilient Agriculture, Indian Agricultural Research Institute, IARI PUSA, New Delhi 110 012, India.
[30]Landscape & Water Sciences, Department of Economic Development, Jobs, Transport and Resources, Horsham 3400, Australia.
[31]Natural Resources Institute Finland (Luke), Latokartanonkaari 9, 00790 Helsinki, Finland
[32]INRA, US1116 AgroClim, 84 914 Avignon, France.
[33]NASA Goddard Institute for Space Studies, New York, NY 10025, USA.
[34]Computational and Systems Biology Department, Rothamsted Research, Harpenden, Herts, AL5 2JQ, UK.
[35]Biological Systems Engineering, Washington State University, Pullman, WA 99164-6120, USA.
[36]PPPS and WSG & CALM, Wageningen University, 6700AA Wageningen, The Netherlands.
[37]Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Science, Beijing 100101, China.
[38]CSIRO Agriculture, Brisbane, Queensland 4102, Australia.
[39]INRA, UMR 1248 Agrosystèmes et développement territorial (AGIR), 31 326 Castanet-Tolosan, France.

*These authors contributed equally to this work.
||Authors from P.K.A. to Y.Z. are listed in alphabetical order.
†Present address: Institute of Landscape Systems Analysis, Leibniz Centre for Agricultural Landscape Research, 15374 Müncheberg, Germany.
‡Present address: European Commission Joint Research Centre, 21 027 Ispra, Italy.
¶Formerly: Natural Ressources Institute Finland (Luke), 00790 Helsinki, Finland.
§Present address: Department of Plant and Soil Sciences, Oklahoma State University, Stillwater, OK, 74078-6028, USA.
¥Present address: James Hutton Institute, Invergowrie, Dundee, DD2 5DA, Scotland, UK.
ɸPresent address: Institute for Sustainable Food Systems, University of Florida, Gainesville, FL 32611, USA.
€Present address: Institute of Future Environment, Queensland University of Technology, Brisbane, QLD 4001, Australia.
¢Present address: CSIRO, Queensland Bioscience Precinct St Lucia, St Lucia, Qld 4067, Australia.
£The views expressed in this paper are the views of the authors and do not necessarily represent the views of the organization or institution, with which he is currently affiliated.

Increasing the accuracy of crop productivity estimates is a key element in planning adaptation strategies to ensure global food security under climate change. Process-based crop models are effective means to project climate impact on crop yield, but have large uncertainty in yield simulations. Here we show that variations in the mathematical functions currently used to simulate temperature responses of physiological processes in 29 wheat models account for >50% of uncertainty in simulated grain yields for mean growing season temperatures from 14°C to 33°C. We derived a set of new temperature response functions that when substituted in four wheat models reduced the error in grain yield simulations across 7 global sites with different temperature regimes by 19% to 50% (42% average). We anticipate the improved temperature responses to be a key step to improve modeling of crops under rising temperature and climate change, leading to higher skill of crop yield projections.

Process-based modeling of crop growth is an effective way of representing how crop genotype, environment and management interactions affect crop production to aid tactical and strategic decision making[1]. Process-based crop models are increasingly used to project impact of climate change on crop yield[2]. However, current models produce different results, creating large uncertainty in crop yield simulations[3]. A model inter-comparison study within the Agricultural Model Inter-comparison and Improvement Project (AgMIP)[4] of 29 widely used wheat models against field experimental data revealed that there is more uncertainty in simulating grain yields from the different models than from 16 different climate change scenarios[3]. The greatest uncertainty was in modeling crop responses to temperature[3,5]. Similar results were found with rice and maize crops[6,7]. Such uncertainty should be reduced before informing decision-making in agriculture and government policy. Here we show contrasting differences in temperature response functions of key physiological processes adopted in the 29 crop models. We reveal opportunities for improving simulation of temperature response in crop models to reduce the uncertainty in yield simulations.

We aim to reassess the scientific assumptions underlying model algorithms and parameterization describing temperature-sensitive physiological processes, using wheat, one of the most important staple crops globally, as an example. We hypothesized that: 1) the difference among models in assumed temperature responses is the largest source of the uncertainty in simulated yields; and 2) the uncertainty in the multi-model ensemble results can be reduced by improving the science for modeling temperature response of physiological processes.

Temperature affects crop performance primarily through its impact on 1) the rate of phenological development from seed germination to crop maturity, including the fulfillment of cold requirement (vernalization); 2) the initiation and expansion of plant organs; 3) photosynthesis and respiration, considered either separately or combined as net biomass growth simulated using radiation use efficiency (RUE)[8]; and 4) the senescence, sterility or abortion of plant organs. All 29 models simulate these processes, except for sterility and abortion, in response to temperature change.

Here, we compare the temperature functions of these four categories of physiological processes built into the 29 wheat models and identify the representative response types. We analyze how different temperature response functions affected simulations of wheat growth compared to observations in a field experiment[8-10], in which well-fertilized and irrigated wheat grew under contrasting sowing dates and temperature environments (Hot Serial Cereal [HSC] experiment). We further evaluate the impact of the different response types by implementing them in two models (APSIM and SiriusQuality) and analyzing their results against the HSC data and an additional global dataset from the International Heat Stress Genotpye Experiment (IHSGE)[8] carried out by the International Maize and Wheat Improvement Center (CIMMYT). More importantly, we derive, based on newest knowledge and data, a set of new temperature response functions for the key physiological processes of wheat, and demonstrate that when substituted in four wheat models the new functions reduced the error in grain yield simulations across seven global sites with different temperature regimes covered by the IHSGE data.

Results

Contrasting temperature functions in 29 models.

A wide range of temperature responses was observed in the 29 models (Supplementary Table 1 & 2) which we grouped into four major types (type 1-4) according to how phenological development and biomass growth (RUE) are treated (Fig. 1 and Supplementary Table 3), i.e. whether increasing or decreasing slopes are linear or curvilinear, whether base (T_{min}), optimum (T_{opt}) or maximum (T_{max}) temperatures are defined, and whether T_{opt} is a range or a point. The simplest type is a linear increase in developmental rate with temperature from a base temperature (T_{min}) around 0°C assuming no temperature optimum (T_{opt}) or maximum (T_{max}) (type 1 Phenology, Fig. 1a), and a linear decline of biomass growth rate above a certain temperature assuming no T_{min} (type 1 Biomass, Fig. 1b). For both processes, the second type defines both T_{min} and T_{opt}, but assumes no T_{max}, thus simulating an increasing rate with temperature below T_{opt} and a constant maximum rate above T_{opt}, respectively (type 2, Fig. 1, c and d). Most models define the three cardinal temperatures, simulating an increasing rate with temperature from T_{min} to T_{opt} and a decreasing rate from T_{opt} to T_{max} (Fig.1, e, f, g, and h). Some of the models in this category define T_{opt} as a range (type 3, Fig. 1, e and f), while the rest define it as a single value (type 4, Fig. 1, g and h). Some models implement linear responses to temperature between the cardinal temperatures, the others curvilinear.

For both phenology and biomass growth, most models agree on a T_{opt} when the rate is maximum (Fig. 1), except for models that lack a T_{opt} (Fig. 1a). At temperatures lower or higher than T_{opt}, the uncertainty in the simulation of phenological development and biomass increases, particularly at higher temperatures. Response types for photosynthesis were consistent, but different cardinal temperatures were used introducing uncertainty (Supplementary Fig. 1, a and b). The simulated temperature responses of respiration differ widely from each other (Supplementary Fig. 1, c and d). When such estimates of respiration and photosynthesis are combined to simulate growth, any uncertainty is compounded at high temperatures. For leaf growth and senescence, contrasting temperature responses were deployed, with much greater uncertainty at temperatures above 25-30°C (Supplementary Fig. 1, e, f, g, and h). For grain growth, the differences in temperature responses are even greater, generating increased uncertainty above 24°C (Supplementary Fig. 1, i and j).

Model performance against HSC data.

Simulation results of the 29 models against the HSC experiment were analyzed by grouping all the models based on the four temperature response types and cardinal temperatures deployed for simulating phenology and biomass growth. The results were standardized at 20°C to remove any systematic bias and compare their response to temperature (Fig. 2). For phenology, the models agreed most closely with each other at a mean growing season temperature around 20°C and matched the observed anthesis and maturity dates well (Fig. 2, a to h). At lower and higher temperatures, the simulated results departed from each other and did not match the observed dates. Three type 4 response models (with three cardinal temperatures, Fig. 1g) with low T_{opt} and T_{max} severely underestimated the pre-flowering development rate at temperatures above 25°C and thus predicted durations longer than were observed (Fig. 2d). For post-flowering development, 20 out of the 29 models predicted the physiological maturity to be later than was observed at temperatures above 25°C (Fig. 2, e to h), particularly the models that have a T_{max} around 35°C (Fig. 2h).

For total above ground biomass and grain yield, the models with type 2 response for biomass growth (no reduction at higher temperatures) tended to overestimate biomass at high temperatures (Fig. 2j). For type 3 (with an optimal temperature range, Fig. 2k) and type 4 (Fig. 2l) responses, the models that have a higher T_{opt} and T_{max} for either RUE (Fig. 1, f and h) or photosynthesis (Supplementary Fig. 1a) also overestimated biomass at temperatures above 25°C (Fig. 2, k and l). The simulated responses for grain yield for the HSC experiment varied in a similar way to those for biomass (Fig. 2, m to p). These findings indicate that improved modeling of temperature responses of phenological development, biomass growth (RUE), photosynthesis and respiration rates is necessary to reduce uncertainty in simulation of grain yield.

Impact of temperature response functions.

While the impact of the temperature functions in different models may be compounded by interactions with other simulated processes, we further evaluated the impact of the different temperature response types (Supplementary Table 3) by implementing 20 combinations of temperature response types in the APSIM and SiriusQuality models to simulate the HSC data and the additional IHSGE data from CIMMYT[8,11,12]. This change caused the two models to predict different grain yields as a result of differences in simulated growth duration, leaf area index, and biomass (Fig. 3). Differences in simulated grain yield were greater than 100%, particularly at low and high temperatures (Fig. 3). The range of simulated grain yield caused by different combinations of temperature response functions in APSIM and SiriusQuality was on average 52% (65%) and 64% (78%) of the uncertainty of the whole ensemble of 29 models for the HSC (IHSGE) data, respectively, highlighting the significant impact of temperature response functions alone on simulated wheat growth in the absence of water and nutrient stresses.

New temperature response functions.

A recent synthesis of available data on phenological development and tissue expansion indicated that rates of pre-anthesis phenological development, tissue expansion, and cell division of crop plants all followed a common Arrhenius-type response curve, and for wheat the response curve has a T_{min} of 0°C, T_{opt} of 27.7°C, and T_{max} of 40°C[13,14]. We used this information to derive and unify the modeling of the temperature response for wheat phenological development and initiation and expansion of leaves, nodes, tillers, stem, grain, and roots using a non-linear function (f(T)[15] (Fig. 4, a and equation 1). If such a temperature response represents the crop's development of sink capacity[13], leaf photosynthesis under current CO_2 levels, typical radiation and stress-free conditions should closely follow this response, with T_{opt} around 27.7°C (Fig. 4c), although the T_{opt} of C_3 crops such as wheat may increase under higher CO_2 concentrations and light intensities when photorespiration is suppressed[16].

Data on Q_{10} (the factor by which the rate of a process increases when temperature is raised by 10°C) for various species living in a wide temperature range[17] enabled us to derive cardinal temperatures for respiration using the f(T) equation (Fig. 4c). This new function can accurately simulate the decline in Q_{10} with increasing temperature (Fig. 5), and is similar to that estimated for Eucalyptus pauciflora[18]. This clearly demonstrates the need to replace the traditional constant Q_{10} approach to better quantify the temperature response of respiration. The rates of post-anthesis development calculated with data from experiments in outdoor climate chambers[19] and the HSC experiment, together with the f(T) equation, enabled derivation of the cardinal temperatures of post-anthesis development (Fig. 4b). The rates of post-anthesis development increases with temperature up to 25-30°C[20,21].

We used the derived response functions for photosynthesis and respiration combined with the SPASS canopy photosynthesis and growth model[22] to generate the temperature response for RUE (Supplementary Fig. 2a, Fig. 4d). The emergent response showed a T_{opt} of 20°C, T_{min} of -1°C and T_{max} of 35°C under moderate to high radiation, but T_{opt} shifted towards lower temperatures under low radiation (data not shown), giving a wider T_{opt} range (Supplementary Fig. 2a). The same f(T) equation with these derived cardinal temperatures for RUE (Fig. 4d) is able to explain 99% of the variance of the emergent responses generated from the SPASS model (Supplementary Fig. 2b).

The derived temperature response functions captured real responses well, compared to the pre-anthesis developmental rates reported[13] and calculated from the HSC experimental data (Fig. 4a), post-anthesis developmental rates estimated from an additional data set for a winter wheat cultivar grown in outdoor climate chambers[23] (Fig. 4b), and measured leaf photosynthesis rates[24] (Fig. 4c). Pooling all data, the derived response functions explained 84% (for post-anthesis development) to 95% (for seedling elongation) of the variation in the rates calculated from measured data (Supplementary Fig. 3). The derived temperature function for RUE (Fig. 4d) matched the response of maximum net biomass growth rates calculated from the HSC, and that of the maximum RUE calculated from LAI, biomass and radiation interception for two additional data sets for winter wheat grown in the field in the North China Plain (NCP)[25] and in outdoor climate chambers[19]. A comparison of the net biomass growth rate and RUE for the NCP and outdoor climate chamber experiments (Supplementary Fig. 4) demonstrated that under the current CO_2 level, RUE for biomass growth under conditions free of other stresses follows the temperature response shown in Fig. 4d, representing the upper boundary of the calculated RUE across a wide temperature range, and is consistent with previous studies[24]. Except for the responses of daily biomass growth and RUE where daily average temperatures are used, use of sub-daily temperatures and canopy temperatures may further improve the simulated response.

Improvement in wheat yield simulations.

Implementation of the derived temperature response functions in APSIM and SiriusQuality improved the simulation of wheat phenological development, biomass growth and grain yield across growing temperatures from 15°C to 32°C compared with data from both HSC experiment and the independent IHSGE global experiment (Fig. 3). For HSC, only the post-anthesis development rates were used to derive f(T) so that data can be considered as semi-independent. Compared with the original models, the root mean squared relative error (RMSRE) of the models for grain yield with the derived temperature responses was reduced by 58% (from 58% to 24%) and 53% (from 53% to 25%) for APSIM and SiriusQuality, respectively against the HSC data. The error reduction for the IHSGE data set was 60% (from 100% to 39%) and 39% (from 31% to 19%) for APSIM and SiriusQuality, respectively.

The improved temperature functions were tested further using two additional models (SALUS and WheatGrow) with the multi-environment IHSGE experimental data

(Table 1). Improvements in simulating total biomass and grain yields were demonstrated in all the four models, with a reduction in RMSE by 28-60% for biomass and 19-59% for grain yield. Less improvement was achieved for modelling phenological development for both models, possibly due to an over-fitting of the original models as phenological data were provided to modelers and models were not fully recalibrated after the implementation of the improved equations. The four improved models had a larger modeling efficiency for both total biomass and grain yield (Table 1), indicating that they better captured the variations of these variables to temperature. We conclude that the common equation f(T) with different parameters for different processes is able to simulate the temperature responses of major physiological processes in wheat and may be potentially applied to other crops to increase certainty in simulating crop yield under climate change[13,14].

Discussion

With the increased applications of process-based crop models to address genotype × environment × management interactions as they impact on yield under climate change, the science underpinning a model for simulation of crop growth processes and yield needs to be critically examined to ensure high scientific rigor and simulation certainty. Our analyses revealed contrasting differences in the type of mathematical equations used to simulate temperature responses of the key physiological processes of wheat. Such differences are a major cause for large uncertainty in simulated wheat yields across different temperature environments. They also reflect the insufficient understanding of how key physiological processes respond to temperature at the time when the models were originally developed, many of which were only based on limited data and local conditions. We demonstrated that by updating the temperature response functions based on newest science and data, crop models can better capture the impact of temperature change on growth processes and gain yield, unveiling a major step to improve modeling of crops under rising temperature and climate change, leading to higher skill of crop yield projections.

The Agricultural Model Inter-comparison and Improvement Project (AgMIP) has enabled a worldwide comparison of agricultural models against global datasets. The inter-comparison of 29 wheat models showed that uncertainty in simulated wheat yield from different models increases with rising temperature, which provides the background and forms the basis for our current study. Previous results from a multi-model ensemble approach for wheat[3,5], rice, maize and potato crops[6,7,26] indicated that the mean simulated crop yield of a multi-model ensemble agreed reasonably well with observations, pointing to the use of a multi-model ensemble approach as an effective way of quantifying and reducing uncertainty in crop yield projections under climate change. However, such agreement will ultimately depend on how the response functions for all major physiological processes compare among the models and how closely they are to the 'true' response to environmental variables like temperature. Although the multi-model ensemble approach provides one useful way of uncertainty quantification, it is expensive and difficult to apply in terms of labor, timing and expertise. In addition, the ensemble approach itself does not necessarily lead to improvement in process understanding, unless a further step is taken to increase the rigor of science underpinning the process submodules by improving algorithms in comparison to data, as demonstrated here.

Further analysis of our newly derived response functions reveal that the median responses from all the 29 models closely matched the derived temperature responses for pre-anthesis phenological development from 0°C to 30°C, and for biomass growth rates, RUE and respiration in the range of 0°C to 35°C. However, for post-flowering phenological development, the ensemble median only matched the derived responses up to 25°C, while the median model photosynthesis response matched the derived temperature response of RUE rather than that of photosynthesis (Fig. 4e). The deviations of temperature response functions for various processes in individual models from the newly derived functions based on experimental data imply that there is no guarantee for the multi-model ensemble median or mean to provide the best yield predictor, particularly at high temperatures. Our results highlight the importance of careful ex-ante screening and evaluation of the individual models for their robustness to simulate temperature responses before they are selected in a multi-model ensemble for the purpose of reducing uncertainty in assessment of climate change impact.

Our analyses identified several key knowledge gaps. Very limited data are available to quantify wheat response to extreme temperatures, at both low and high temperature ranges. Further research is needed for the post-anthesis development rate under high temperatures, where models disagree with each other and only few data are available. The models that simulate photosynthesis tend to underestimate T_{opt} for this process and thus need to be reparametrized. There is still a lack of measurement data to quantify how net biomass growth rate responds to temperature, and to verify simulated RUE response to temperature. More generally, variations in vapor pressure deficit (VPD) among environments could introduce noise in the temperature response due to differences in evaporative cooling confounding the association between air and actual plant temperature and thereby reduce the certainty of prediction. Pollination, sterility or abortion of plant organs as affected by abnormal temperatures have rarely been simulated, but can become important under rising temperature, thus needs more attention. While our current analyses focus only on temperature, interactions with other climate drivers will also need to be addressed, for example, interactions with photoperiod on flowering, with radiation on growth rate, with CO_2 concentration change under stressed and non-stressed conditions.

Methods

Inter-comparison of temperature responses in wheat crop models. Twenty-nine physiologically based wheat crop models previously used in the AgMIP-Wheat project[8] (Supplementary Table 1, Supplementary Dataset) were compared in terms of how the key temperature-responsive physiological processes are simulated. The different approaches used in the models are summarized in Supplementary Table 2 and Extended Database 1. The algorithms used in these models were extracted and the temperature response equations for key developmental and growth processes were categorized based on whether the cardinal temperatures (i.e., minimum T_{opt}, optimum T_{opt}, and maximum T_{max}) are defined and if so how. For phenology and biomass four temperature types were identified (Fig 1, Supplementary Fig. 1, and Supplementary Table 3).

Comparison of model performance against data from the Hot-Serial-Cereal experiment. The 29 wheat models were tested against field data from a Hot-Serial-Cereal (HSC) experiment in which the spring wheat cultivar Yecora Rojo was grown with different sowing times and artificial infrared heat treatments under field conditions at Maricopa, AZ, USA (33.07° N, 111.97° W, 361 m a.s.l.)[9,27]. Yecora Rojo is of short stature, requires little to no vernalization, has a low photoperiod sensitivity, and matures early[28]. All crops were well watered and fertilized with temperature being the most variable factor.

The inter-comparison of model performance was part of the AgMIP-Wheat project, with four steps and different levels of available information for model calibration[8]. The results used in this study (Fig. 2, Fig. 3) were simulation results from all models that were calibrated against observed phenology (flowering and maturity dates) from all treatments, together with the in-season and final, total above ground, leaf, stem, and grain dry mass and nitrogen, and leaf area index from the highest yielding treatment, i.e., simulation step D “Blind test with calibrated highest yield”[8].

The HSC data set was also used to assess the uncertainty in the multi-model ensemble due to different types of temperature response functions for phenological development, leaf area index, biomass growth and grain yield (see below).

Evaluation of wheat models against global multi-site experiments. The 29 wheat models were also evaluated against data from the International Heat Stress Genotype Experiments(IHSGE) carried out by CIMMYT (therefore referred to as IHSGE dataset) that had seven temperature environments, including time-of-sowing treatments[11,12], in the absence of water and nutrient stresses and free of pest and diseases. The IHSGE experiments included two spring wheat cultivars (Bacanora 88 and Nesser) grown during the 1990-1991, and 1991-1992 winter cropping cycles at hot, irrigated, and low latitude sites in Mexico (Ciudad Obregon, 27.34° N, 109.92° W, 38 m a.s.l.; and Tlatizapan, 19.69° N, 99.13° W, 940 m a.s.l.), Egypt (Aswan, 24.1° N, 32.9° E, 200 m a.s.l.), India (Dharwar, 15.49° N, 74.98° E, 940 m a.s.l.), the Sudan (Wad Medani, 14.40° N, 33.49° E, 411 m a.s.l.), Bangladesh (Dinajpur, 25.65° N, 88.68° E, 29 m a.s.l.), and Brazil (Londrina, 23.34° S, 51.16° W, 540 m a.s.l.)[11,12,29]. Experiments in Mexico included normal (December) and late (March) sowing dates. Bacanora 88 has moderate vernalization requirement and low photoperiod sensitivity and Nesser has low to no vernalization requirement and photoperiod sensitivity. All experiments were well watered and fertilized with temperature being the most important variable. Variables measured in the experiment included plants m^{-2}, total above ground biomass at 50% anthesis, days to 50% anthesis, days to physiological maturity, final total above ground biomass, grain yield, spikes m^{-2}, grains $spike^{-1}$, and average single grain mass at maturity.

Model inter-comparison was carried out using standardized protocols and one step of calibration[8]. These experimental data were not publicly available and were therefore used in a blind test. Sowing dates, anthesis and maturity dates, soil type characteristics and weather data for all sites, years, and cultivars were supplied to the modellers. Crop growth data were supplied only for one site (at Obregon) in one year, all other crop growth data were hold back and not supplied to modelers. The IHSGE dataset was also used to assess the uncertainty of the multi-model ensemble due to different types of temperature response functions for phenological development, leaf area index, biomass growth, and grain yield (see below). None of these data were used to derive the improved temperature response functions.

Evaluation of the impact of various temperature response functions on simulation results. In order to demonstrate the impact of the temperature response types used in different wheat crop models on simulated phenology, total above biomass and grain yield, the four major types of temperature responses summarized from the models (Supplementary Table 3) were implemented in the APSIM and SiriusQuality models. These two models were chosen because they were built with different types of temperature response functions (Supplementary Table 3) and use different approaches to simulate phenology (progress to flowering by calculating the duration of phases between significant events on the shoot apex vs. tracks development through leaf appearance, using the prediction of final main stem leaf number), canopy expansion (branching vs. individual phytomer-based approaches), and biomass growth (radiation use efficiency of whole canopy vs. individual canopy layers). For phenology, we also separated the response type 4 into linear and curvilinear responses, resulting in a total of 20 temperature (4 × 5) response type combinations for models using radiation use efficiency (RUE; Supplementary Table 3). The two modified models were executed against the HSC and IHSGE experimental data. For any given observed grain yield, the simulated yield ranges from the multi-model ensemble (of the 29 wheat models), the APSIM and SiriusQuality models (each with the 20 combinations of temperature response functions), were calculated. The ratios of the simulated ranges of the APSIM and SiriusQuality with the 20 combinations of temperature response functions to those of the multi-model ensemble were used to estimate how much variations in the multi-model ensemble ranges were explained by each of the models together with the variations in temperature functions.

New temperature response functions of wheat physiological processes derived based on data. The Wang-Engel (WE) curvilinear temperature response function used to model wheat phenology[15] in the SPASS (Soil Plant Atmosphere Systems Simulation)-Wheat model[30] was found to be accurate and flexible in simulating the temperature responses of wheat plants[31,32]. It has been successfully applied in modeling leaf development and phenology of wheat [31,32], maize[33], rice[34], and potato crops[35].

The WE temperature function constructs a curvilinear response based on T_{min}, T_{opt}, and T_{max} of the simulated process. These three cardinal temperatures determine the shape of the response curve, so they have clear biological meanings. Once the cardinal temperatures are known, no extra parameters are needed in the model. It simulates the effect [0-1] of temperature between T_{min} and T_{max} as:

$$f(T) = \left[\frac{2(T - T_{min})^{\alpha}(T_{opt} - T_{min})^{\alpha} - (T - T_{min})^{2\alpha}}{(T_{opt} - T_{min})^{2\alpha}} \right]^{\beta}; \ \alpha = \frac{\ln 2}{\ln\left(\frac{T_{max} - T_{min}}{T_{opt} - T_{min}}\right)}, \ \beta = 0\sim1 \qquad (1)$$

An extra shape factor β was added here in Equation (1) to account for temperature responses with more extended T_{opt} (e.g. for RUE at low radiation). For all processes β = 1.0 was used to describe temperature responses, except for RUE where $\beta = 0.8$ was used to reflect the different shape of the RUE response curve compared to other physiological processes.

The cardinal temperatures derived for using Equation (1) to simulate temperature responses of various processes are given in Fig 4. For phenological development, the cardinal temperatures were derived from published data on seedling elongation and pre-anthesis development[13], and post-anthesis development[10,23] (see below). For photosynthesis under current CO_2, the cardinal temperatures of pre-anthesis phenological development were used assuming it mimics the development of sink capacity. For respiration rate, Equation (1) with $\beta = 1.0$ was used to derive the average Q_{10} (the factor by which the respiration rate increases when temperature is raised by 10°C) of respiration rate at different temperatures from 5 to 45°C with 5°C interval. A genetic algorithm was applied to optimize the three cardinal temperatures (T_{min}, T_{opt}, and T_{max}) to match the derived average Q_{10} to the Q_{10} estimated at the corresponding temperatures known from the literature[17] (Fig. 5). Finally, for RUE the cardinal temperatures were derived from simulation results using the SPASS canopy photosynthesis and growth model, together with the derived temperature functions for photosynthesis and respiration (see below). All rates were normalized at 20°C.

Calculation of pre- and post-anthesis development rates from data. Pre-anthesis development rates were calculated from the HSC experiment. The rates of leaf emergence were estimated as the slope of the decimal number of emerged leaves (Haun index[36]) measured at least twice a week against days from seedling emergence[37]. The rate of development towards anthesis was calculated as the reciprocal of the duration from emergence to anthesis. The rates of seedling elongation for seven spring wheat cultivars grown in growth chambers with different temperature were also obtained from a recent data synthesis [13].

Post-anthesis rate of development was calculated as the reciprocal of the time from anthesis to physiological maturity from the HSC data[10] and experiments carried at INRA Clermont-Ferrand, France (44.78° N, 3.17° E, 329 m a.s.l.) with the winter wheat cultivar Thésée grown during the 1993-1994, and 1997-1998 winter cropping cycles in outdoor climate chambers under well-watered and fertilized conditions with post-anthesis mean daily temperature ranging from 12.6 to 24.7°C[23]. In the HSC experiment, physiological maturity was judged when endosperm of grains becomes firm and almost dry. In the INRA experiments, physiological maturity was calculated as the time when 95% of final grain dry mass was reached by fitting a 3-parameter logistic function equation to grain dry mass data plotted against the number of days after anthesis[23].

The calculated post-anthesis rate of development from the HSC data was the only data used for derivation of temperature response functions shown in Fig 4. No data from the IHSGE dataset were used in the derivation of temperature functions. Therefore, for model testing, the IHSGE dataset is fully independent data, while the HSC dataset is semi-independent.

Derivation of the emergent temperature response for radiation use efficiency (RUE) using a canopy photosynthesis and growth model. A simplified version of the canopy photosynthesis and growth sub-models in the SPASS-Wheat model[30], together with the derived temperature response functions for photosynthesis and respiration rates (Fig 4c), were used to calculate the net biomass growth rate of a wheat canopy and derive the cardinal temperatures and shape parameter of the RUE temperature response function (Supplementary Fig. 2). The model integrates leaf level photosynthesis rate to canopy level. It also calculates the growth and maintenance respiration, then the net assimilation and net biomass growth. All the parameter values used in the simulations are given in Supplementary Table 4.

We assumed a wheat canopy at an early developmental stage with a leaf area index (LAI) of 3 m^2 m^{-2} and a total above ground biomass of 3 t ha^{-1}. For any new growth, 20% of assimilate would be partitioned to roots, and 80% to the above ground parts. In the simulations, we used 47 years (1957-2003) of daily climate data from Birchip in Victoria of Australia to simulate the daily RUE of the wheat canopy in the absence of water and nutrient stresses. This gave us a daily global radiation range from 10 to 32 MJ d^{-1} and a daily mean temperature range of 3.6 to 36°C. We also executed the model for an extra range of daily mean temperature from -5 to 5°C to generate the daily net above ground biomass growth rate. RUE was calculated for different daily temperatures as the net above ground biomass growth rate divided by the radiation intercepted by the canopy.

Calculation of net biomass growth rate and radiation use efficiency under different temperatures. Net biomass growth rate was calculated from the HSC data as the ratio of total above ground biomass at maturity divided by the number of days from crop emergence to physiological maturity. Measurement data on dynamics of LAI and total above ground biomass from the INRA experiments described above [19] and from five experiments where the winter wheat cultivars SJZ8 and SJZ15 were grown during the 2004-2005, 2005-2006, 2006-2007, and 2009-2010 winter cropping cycles at Wuqiao, North China Plain (NCP, 37.41° N, 116.37° E, 20 m a.s.l.) with ample water and nitrogen supply [25], were used to calculate RUE under different temperatures.

In the INRA experiments, LAI and total above ground biomass were measured every 4 to 8 days starting at anthesis. Only dates when LAI was higher than 2.5 m^2 m^{-2} were used (i.e. before the onset of the phase of rapid canopy senescence), leaving measurements from five to six dates with which to calculate the net biomass growth rate and RUE. Daily radiation interception was calculated as total incident radiation times (1-exp(-$K_L \times$ LAI)), where K_L (0.7 m^2 ground m^{-2}

green leaf) is the light extinction coefficient. RUE was calculated as the slope of total above ground biomass versus the cumulative radiation interception and the average net biomass growth rate was calculated as the slope of total aboveground biomass versus the number of days after anthesis.

In the NCP experiments, LAI and total above ground biomass were measured before wintering, at greening, and at jointing, booting, anthesis and 10 days after flowering and at maturity. Daily increases in LAI were estimated through best fit polynomial equations to the data. Daily radiation interception was calculated as for the INRA experiments but using total incident radiation estimated from sunshine hours. The cumulative radiation interception for each period was calculated as the sum of daily radiation interception. RUE for each period (from jointing onwards) was calculated as the net biomass increase divided by the total radiation interception and the average net biomass growth rate was also calculated for each period (from jointing onwards) as the net biomass increase divided by the total number of days.

Calculation of daily mean temperature. Daily mean air temperature (Tt) in the HSC and IHSGE experiments was calculated as the sum of eight contributions of a cosine variation between daily maximum ($T_{max,daily}$) and minimum ($T_{min,daily}$) daily air temperatures[38]:

$$T_t = \frac{1}{8}\sum_{r=1}^{r=8} T_h(r) \qquad (2)$$

$$T_h(r) = T_{min,daily} + f_r\left(T_{max,daily} - T_{min,daily}\right) \qquad (3)$$

$$f_r = \frac{1}{2}\left(1 + \cos\frac{90}{8}(2r-1)\right) \qquad (4)$$

where T_h (°C) is the calculated 3-hour temperature contribution to estimated daily mean temperature, and r is an index for a particular 3-hour period.

Evaluation of the improved temperature response functions. We tested the performance of the new temperature response functions on how accurately they capture the rates of the phenological development, tissue expansion, photosynthesis, and biomass growth (RUE) measured or derived from experimental data at a range of temperatures. This was done by comparing the rates calculated using the derived functions (Fig. 4) at a given temperature against the corresponding measured rates from the experiments at the same temperature (Supplementary Fig. 3-4). Significance of the relationship was tested and the coefficient of determination (R^2) was used to see how much variation in the measurements could be explained by the new temperature functions.

Evaluation of the improved skills of four wheat models when using the new temperature responses. To test the improvement by using the improved temperature response functions, they were also implemented into the APSIM, SiriusQuality, SALUS and WheatGrow models replacing their original functions. The simulation results were then compared with the measurements (Fig 3, Table 1). These four models were chosen to have good representation of different temperature response functions for phenological development and biomass growth and thus to generalize the improvement in wheat model skills when they use the temperature response function we derived. One of the models (WheatGrow) uses a photosynthesis and respiration approach to model biomass growth, while the other three use a RUE approach.

Many different measures of the discrepancies between simulations and measurements have been proposed[39]. We concentrated on three measures to highlight different aspects of the quality of simulation with the original and improved models. All measures are based on mean squared error (MSE), where the mean is over all measurements of a particular variable. The root mean squared error (RMSE) is the square root of MSE; it has the advantage to express errors in the same units as the variable. RMSE was calculated as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2} \tag{5}$$

where y_i is the observed value of the ith measured treatment, $\hat{y}_i$ is the corresponding simulated value, and N is the total number of treatments.

For comparing very different growth environments likely to give a broad range of crop responses, the relative error can be more meaningful than the absolute error, so the root mean squared relative error (RMSRE) was also calculated because of the very wide range of total above ground biomass and grain yields in both the HSC and IHSGE datasets. RMSRE was calculated as:

$$\mathrm{RMSRE} = 100 \times \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i - \hat{y}_i}{y_i}\right)^2} \tag{6}$$

Finally, the Nash–Sutcliffe model efficiency[40] (EF) is a distance measure that compares model MSE with the MSE of using the average of measured values as an estimator. Therefore, EF is useful for making statements about the skill of a model relative to this simple reference estimator. For a model that simulates perfectly, EF = 1, and for a model that has the same squared error of simulation as the mean of the measurements, EF = 0. EF is positive for a model that has a smaller squared error than the mean of the measurements. EF was calculated as:

$$\mathrm{EF} = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2} \quad (7)$$

where $\bar{y}$ is the average over the y_i.

Data availability. The data extracted from the models to describe their temperature functions are provided in Supplementary_Data_Set_D1 in Excel format. The experimental data used to calibrate and validate the models are available in Harvard Dataverse with the identifiers "doi:10.7910/DVN/1WCFHK"[41] for and HSC data and "doi:10.7910/DVN/ECSFZG"[42] for the IHSGE data.

References

1 Porter, J. R. & Semenov, M. A. Crop responses to climatic variation. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**, 2021-2035 (2005).

2 Liu, B. *et al.* Similar estimates of temperature impacts on global wheat yield by three independent methods. *Nature Clim. Change* **advance online publication** (2016).

3 Asseng, S. *et al.* Uncertainty in simulating wheat yields under climate change. *Nature Climate Change* **3**, 827-832 (2013).

4 Rosenzweig, C. *et al.* The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. *Agricultural and Forest Meteorology* **170**, 166-182 (2013).

5 Rotter, R. P., Carter, T. R., Olesen, J. E. & Porter, J. R. Crop-climate models need an overhaul. *Nature Clim. Change* **1**, 175-177 (2011).

6 Li, T. *et al.* Uncertainties in predicting rice yield by current crop models under a wide range of climatic conditions. *Glob Chang Biol* **21**, 1328-1341 (2015).

7 Bassu, S. *et al.* How do various maize crop models vary in their responses to climate change factors? *Global Change Biology* **20**, 2301-2320 (2014).

8 Asseng, S. *et al.* Rising temperatures reduce global wheat production. *Nature Climate Change* **5**, 143-147 (2015).

9 Wall, G. W., Kimball, B. A., White, J. W. & Ottman, M. J. Gas exchange and water relations of spring wheat under full-season infrared warming. *Global Change Biology* **17**, 2113-2133 (2011).

10 White, J. W., Kimball, B. A., Wall, G. W., Ottman, M. J. & Hunt, L. A. Responses of time of anthesis and maturity to sowing dates and infrared warming in spring wheat. *Field Crops Research* **124**, 213-222 (2011).

11 Reynolds, M., Balota, M., Delgado, M., Amani, I. & Fischer, R. Physiological and morphological traits associated with spring wheat yield under hot, irrigated conditions. *Functional Plant Biology* **21**, 717-730 (1994).

12 Reynolds, M. P. *et al.* The International Heat Stress Genotype Experiment: results from 1990-1992 184-192 (CIMMYT,DF, Mexico, 1994).

13 Parent, B. & Tardieu, F. Temperature responses of developmental processes have not been affected by breeding in different ecological areas for 17 crop species. *New Phytologist* **194**, 760-774 (2012).

14 Parent, B., Turc, O., Gibon, Y., Stitt, M. & Tardieu, F. Modelling temperature-compensated physiological rates, based on the co-ordination of responses to temperature of developmental processes. *Journal of Experimental Botany* **61**, 2057-2069 (2010).

15 Wang, E. & Engel, T. Simulation of phenological development of wheat crops. *Agricultural Systems* **58**, 1-24 (1998).

16 Yin, X. & Struik, P. C. C3 and C4 photosynthesis models: An overview from the perspective of crop modelling. *NJAS - Wageningen Journal of Life Sciences* **57**, 27-38 (2009).

17 Atkin, O. K. & Tjoelker, M. G. Thermal acclimation and the dynamic response of plant respiration to temperature. *Trends in Plant Science* **8**, 343-351 (2003).

18 O'Sullivan, O. S. *et al.* High-resolution temperature responses of leaf respiration in snow gum (*Eucalyptus pauciflora*) reveal high-temperature limits to respiratory function. *Plant, Cell & Environment* **36**, 1268-1284 (2013).

19 Martre, P., Porter, J. R., Jamieson, P. D. & Triboï, E. Modeling grain nitrogen accumulation and protein composition to understand the sink/source regulations of nitrogen remobilization for wheat. *Plant Physiology* **133**, 1959-1967 (2003).

20 Hunt, L. A., Poorten, G. v. d. & Pararajasingham, S. Postanthesis temperature effects on duration and rate of grain filling in some winter and spring wheats. *Canadian Journal of Plant Science* **71**, 609-617 (1991).

21 Sofield, I., Evans, L., Cook, M. & Wardlaw, I. Factors influencing the rate and duration of grain filling in wheat. *Functional Plant Biology* **4**, 785-797 (1977).

22 Wang, E. & Engel, T. SPASS: a generic process-oriented crop model with versatile windows interfaces. *Environmental Modelling & Software* **15**, 179-188 (2000).

23 Triboï, E., Martre, P. & Triboï-Blondel, A.-M. Environmentally-induced changes in protein composition in developing grains of wheat are related to changes in total protein content. *Journal of Experimental Botany* **54**, 1731-1742 (2003).

24 Nagai, T. & Makino, A. Differences between rice and wheat in temperature responses of photosynthesis and plant growth. *Plant and Cell Physiology* **50**, 744-755 (2009).

25 Zhao, Z. *et al.* Accuracy of root modelling and its impact on simulated wheat yield and carbon cycling in soil. *Field Crops Research* **165**, 99-110 (2014).

26 Fleisher, D. H. *et al.* A potato model intercomparison across varying climates and productivity levels. *Global Change Biology* **23**, 1258-1281 (2017).

27 Ottman, M. J., Kimball, B. A., White, J. W. & Wall, G. W. Wheat growth response to increased temperature from varied planting dates and supplemental infrared heating. *Agron. J.* **104**, 7-16 (2012).

28 Qualset, C. O., Vogt, H. E. & Borlaug, N. E. Registration of ‘Yecora Rojo’ wheat. *Crop Science* **25**, 1130 (1985).

29 Reynolds, M. P. in *Wheat in heat stressed environments: irrigated dry areas and rice-wheat farming systems* (eds D.A. Saunders & G.P. Hettel) 184-192 (DF, 1993).

30 Wang, E. & Engel, T. Simulation of growth, water and nitrogen uptake of a wheat crop using the SPASS model. *Environmental Modelling and Software* **17**, 387-402 (2002).

31 Streck, N. A., Weiss, A., Xue, Q. & Stephen Baenziger, P. Improving predictions of developmental stages in winter wheat: A modified Wang and Engel model. *Agricultural and Forest Meteorology* **115**, 139-150 (2003).

32 Xue, Q., Weiss, A. & Baenziger, P. S. Predicting phenological development in winter wheat. *Climate Research* **25**, 243-252 (2004).

33 Streck, N. A., Lago, I., Gabriel, L. F. & Samboranha, F. K. Simulating maize phenology as a function of air temperature with a linear and a nonlinear model. *Pesquisa Agropecuária Brasileira* **43**, 449-455 (2008).

34 Streck, N. A., Bosco, L. C. & Lago, I. Simulating Leaf Appearance in Rice All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any

information storage and retrieval system, without permission in writing from the publisher. *Agron. J.* **100**, 490-501 (2008).

35 Streck, N. A., Uhlmann, L. O., Zanon, A. J. & Bisognin, D. A. Impact of elevated temperature scenarios on potato leaf development. *Engenharia Agricola* **32**, 689-697 (2012).

36 Haun, J. R. Visual Quantification of Wheat Development. *Agronomy Journal* **65**, 116-119 (1973).

37 White, J. W., Kimball, B. A., Wall, G. W. & Ottman, M. J. Cardinal temperatures for wheat leaf appearance as assessed from varied sowing dates and infrared warming. *Field Crops Research* **137**, 213-220 (2012).

38 Weir, A. H., Bragg, P. L., Porter, J. R. & Rayner, J. H. A winter wheat crop simulation model without water or nutrient limitations. *Journal of Agricultural Science* **102**, 371-382 (1984).

39 Wallach, D., Makowski, D., Jones, J. & Brun, F. *Working with Dynamic Crop Models, 2nd Edition: Methods, Tools and Examples for Agriculture and Environment*. (Academic Press, 2013).

40 Nash, J. E. & Sutcliffe, J. V. River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology* **10**, 282-290 (1970).

41 Martre, P. *et al.* The Hot Serial Cereal Experiment for modeling wheat response to temperature: field experiments and AgMIP-Wheat multi-model simulations. *Harvard Dataverse, http://dx.doi.org/10.7910/DVN/ECSFZG* (2017).

42 Martre, P. *et al.* The International Heat Stress Genotype Experiment for modeling wheat response to heat: field experiments and AgMIP-Wheat multi-model simulations. *Harvard Dataverse, http://dx.doi.org/10.7910/DVN/1WCFHK* (2017).

Supplementary Information is available in the online version of the paper.

Author Information:

Correspondence and requests for materials should be addressed to:

Enli.Wang@csiro.au; pierre.martre@ inra.fr

Acknowledgments

The authors thank David Lobell for useful comments on an earlier version of the paper. EW acknowledge the support from the CSIRO project ‘Enhanced modelling of genotype by environment interactions’ and project ‘Advancing crop yield while reducing the use of water and nitrogen’ jointly funded by CSIRO and the Chinese Academy of Sciences (CAS). ZZ received scholarship from the China Scholarship Council through the CSIRO and Chinese Ministry of Education PhD Research Program. PM, AM, DR acknowledge support from the FACCE JPI MACSUR project (031A103B) through the metaprogram Adaptation of Agriculture and Forests to Climate Change (AAFCC) of the French National Institute for Agricultural Research (INRA). AM has received the support of the EU in the framework of the Marie-Curie FP7 COFUND People Programme, through the award of an AgreenSkills fellowship under grant agreement n° PCOFUND-GA-2010-267196. SA and DC acknowledge support provided by the International Food Policy Research Institute (IFPRI), CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), the CGIAR Research Program on Wheat and the Wheat Initiative. CS was funded through USDA National Institute for Food and Agriculture award 32011-68002-30191. CM received financial support from the KULUNDA project (01LL0905L) and the FACCE MACSUR project (031A103B) funded through the German Federal Ministry of Education and Research (BMBF). FE received support from the FACCE MACSUR project (031A103B) funded through the German Federal Ministry of Education and Research (2812ERA115) and EER was funded through the German Federal Ministry of Economic Cooperation and Development (Project: PARI). MJ and JEO were funded through the FACCE MACSUR project by the Danish Strategic Research Council. KCK and CN were funded by the FACCE MACSUR project through the German Federal Ministry of Food and Agriculture (BMEL). FT, TP and RPR received financial support from FACCE MACSUR project funded through the Finnish Ministry of Agriculture and Forestry (MMM); FT was also funded through National Natural Science Foundation of China (No. 41071030). CB was funded through the Helmholtz project `REKLIM-Regional Climate Change: Causes and Effects' Topic 9: `Climate Change and Air Quality'. MPR and PDA received funding from the CGIAR Research Program on Climate Change, Agriculture, and Food Security (CCAFS). GO'L was funded through the Australian Grains Research and Development Corporation and the Department of Environment and Primary Industries Victoria, Australia. RCI was funded by Texas AgriLife Research, Texas A&M University. BB was funded by USDA-NIFA Grant Number: 2015-68007-23133.

Author Contributions:

E.W., P.M., S.A., F.E. motivated the study, E.W. and P.M. designed and coordinated the study, and analyzed data, E.W., P.M., Z.Z., A.M., L.L. and B.B. conducted model improvement simulations, E.W., P.M., S.A., F.E., Z.Z., A.M., R.P.R.,.K.A., P.D.A., J.A., C.B., D.C., A.J.C., G.D.S., J.D., E.F., M.G-V., S.G., G.H., L.A.H., R.C.I., M.J., C.D.J., K.C.K., A-K.K., C.M., L.L., S.N.K., C.N., G.O'L., J.E.O., T.P., E.P., M.P.R., E.E.R., D.R., A.C.R., M.A.S., I.S., C.S., P.S., T.S., I.S., F.T., P.T., K.W., D.W., J.W. and Y.Z. carried out crop model simulations and discussed the results, B.A.K., M.J.O., G.W.W., J.W.W., M.P.R., P.D.A., and Z.W. provided experimental data, E.W. and P.M. analyzed the results and wrote the paper.

Figure legends

Figure 1 | Temperature response functions in 29 wheat simulation models. (**a**, **c**, **e**, **g** and **i**) Phenological development (pre-flowering). (**b**, **d**, **f**, **h** and **j**) Biomass growth (or RUE). (**a** and **b**) Type 1, linear with no optimum or maximum temperature; (**c** and **d**) type 2, linear or curvilinear with an optimum but no maximum temperature; (**e** and **f**) type 3, linear with range of optimal temperatures; (**g** and **h**) type 4, linear or curvilinear with three cardinal temperatures.

(**i** and **j**) Summary of temperature responses of all models, with red lines representing the median and shaded area the 10% and 90% percentiles for the 29 models. In (**a** to **j**) rates are normalized to 20ºC. Models are listed in Supplementary Table 1

Figure 2 | Comparison of multi-model simulations against observations and average growing season temperature. (**a** to **d**) Simulated days from sowing to anthesis. (**e** to **h**) Simulated days from anthesis to maturity. (**i** to **l**) Simulated final total above ground biomass. (**m** to **p**) Simulated final grain yield. The data were standardized to 20°C and plotted against the mean average daily temperature from sowing to anthesis (**a** to **d**), from anthesis to maturity (**e** to **h**), and from sowing to maturity (**i** to **p**). Models were grouped according to their temperature response types for phenological development (**a** to **h**) or biomass growth (**i** to **p**), as defined in Fig. 1. Simulated and experimental data are for the HSC experiment[8]. Symbols with error bars are experimental means ± 1 s.d. for n = 3 independent replicates.

Figure 3 | Uncertainty in simulated wheat responses due to variations in the temperature response functions of phenological development and biomass growth (RUE). Comparisons are between observed and simulated grain yield (**a, b, c, d**), total above ground biomass (**e, f, g h**), crop growth duration (**i, j, k l**), and in season maximum LAI (**m, n, o, p**) for the HSC and IHSGE data sets. Simulations were executed with the wheat models APSIM and SiriusQuality. Red circles show the measurements (mean ± 1 s.d. for n = 3 independent replicates). Green areas show uncertainty in simulated values (10th to 90th percentile range) from the 29 models of the AgMIP-Wheat multi-model ensemble[8]. Blue areas show the range of simulated values when using APSIM or SiriusQuality combined with the 20 combinations of the four or five types of response functions for phenological development and biomass growth, respectively, using the cardinal temperatures reported in Supplementary Table 3. Dashed black lines show the simulated values by the original APSIM and SiriusQuality models. Solid black lines show the simulated values by APSIM or SiriusQuality with the improved temperature response functions for phenological development and biomass growth.

Figure 4 | Derived temperature responses of various physiological processes. The relative rates of pre- (**a**) and post- (**b**) anthesis development, photosynthesis and respiration (**c**), and biomass growth or RUE (**d**) calculated with data from the literature (symbols) were compared with those estimated using the derived temperature response functions (solid lines). In (**c**) a Q_{10} value of 2 was used for response shown with the dashed line. In (**d**) daily RUE (light blue circles) calculated with the SPASS photosynthesis and plant growth model using daily weather data covering temperature range of -5 to 36ºC. The numbers in the brackets in the legends for the response lines indicate the minimum (T_{min}), optimum (T_{opt}) and maximum (T_{max}) temperatures. The numbers in the brackets in the legends for the data symbols indicate the literature reference source of data. In (**e**), the derived responses (lines) were compared with the medians of the temperature responses calculated from all 29 models (symbols). All data were normalized at 20°C and all curves were generated using the f(T) function equation[15] and the cardinal temperatures shown. For all processes $\beta = 1.0$ except for RUE where $\beta = 0.8$.

Figure 5 | Comparison of Q_{10} for respiration derived from the temperature response function in Figure 4C to the temperature dependence of the Q_{10} of foliar respiration rates[17]. Closed symbols are mean Q_{10} of foliar respiration rate of species of arctic (circles, 49 species), boreal (triangles, 24 species), temperate (squares, 50 species), and tropical (diamonds, 3 species) biomes taken from literature[17]. Black dotted lines indicate ± 1 s.d. of all observations across biomes[17]. A single linear regression was fitted to all experimental data (solid black line). The Q_{10} of the respiration rate derived using the non-linear function equation f(T) (equation 1), together with parameters in Figure 4C, is shown (thick blue line). Data are reproduced with permission[17].

Table 1 | Model improvement statistics for simulation of days to maturity, above ground biomass, grain yield and grain number in the independent IHSGE data after implementation of the new temperature response functions of phenological development and biomass growth (RUE) in four wheat models: APSIM, SiriusQuality, SALUS, and WheatGrow

	Grain yield		Total above ground biomass		Days to maturity		Grain number	
Model	Original model	Improved model	Original model	Improved model	Original model	Improved model	Original model	Improved model
Root mean squared error (RMSE)								
	(t ha^{-1})		(t ha^{-1})		(days)		(grain m^{-2})	
APSIM	2.99	1.23	5.91	2.38	12.3	8.3	4647	3732
SiriusQuality	1.05	0.67	2.89	1.84	11.1	11.8	4046	2886
Salus	2.00	0.88	2.56	1.85	10.1	10.7	-	-
WheatGrow	2.43	1.98	5.47	2.95	1.4	3.6	-	-
Modeling efficiency (EF)								
	(-)		(-)		(-)		(-)	
APSIM	-1.91	-0.09	-1.53	0.32	-0.10	0.62	-1.63	-0.78
SiriusQuality	-0.02	0.66	-0.14	0.46	0.32	0.41	-1.52	-0.06
Salus	0.05	0.56	0.53	0.63	0.37	0.62	-	-
WheatGrow	-1.73	-0.58	-1.48	-0.71	0.99	0.93	-	-