Extending a first-principles primary production model to predict wheat yields

Shengchao Qiao[1,2], Han Wang[1,2,*], I. Colin Prentice[1,3,4], Sandy P. Harrison[1,5]

[1]Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 100084, China;

[2]Joint Center for Global Change Studies (JCGCS), Beijing 100875, China;

[3]AXA Chair of Biosphere and Climate Impacts, Department of Life Sciences, Imperial College London, Silwood Park Campus, Buckhurst Road, Ascot, SL5 7PY, UK;

[4]Department of Biological Sciences, Macquarie University, North Ryde, NSW 2109, Australia;

[5]School of Archaeology, Geography and Environmental Sciences (SAGES), University of Reading, Reading, RG6 6AH, UK.

[*]Corresponding author: Han Wang (wanghan_sci@yahoo.com)

Data statement

The climate and flux data from WeiShan can be obtained by contacting Huimin Lei (leihm@tsinghua.edu.cn). The flux data for YuCheng and the climate, LAI, and crop data used in this manuscript are publicly available from the National Ecosystem Research Network of China, CNERN (http://www.cnern.org.cn/). All climate data driving the PC model runs for the future, and the model outputs of LPJmL, EPEIC and GEPIC, are publicly available from Inter-Sectoral Impact Model Intercomparison Project-2b (ISIMIP2b: https://www.isimip.org/protocol/#isimip2b/). The PC model code will be available from Mendeley data.

Abstract

Climate exerts a major influence on crop development and yield. Despite extensive modelling efforts, there is still considerable uncertainty about the consequences of a changing climate for the yields of major crops. Existing crop models are complex and rely on many assumptions and parameters, motivating a quest for more parsimonious models with stronger theoretical and empirical foundations. This paper presents a prototype of such a model for wheat, informed by measurements of gross primary production (GPP), biomass and yield at research sites across the wheat-growing regions of China. First, GPP was predicted using a recently developed first-principles model driven only by climate, carbon dioxide (CO_2) concentration, and light absorbed by leaves. Modelled GPP was shown to agree well with eddy-covariance measurements. Second, the data were used to show that above-ground biomass (AB) is proportional to time-integrated GPP, and that grain yield shows a saturating relationship with AB. Simple empirical equations based on these findings were combined with modelled GPP to predict yield, including propagated errors due to parameter uncertainty in both the GPP model and the empirical equations. The resulting 'hybrid' model, applied in a variety of climates, successfully predicted measured interannual variations in AB and yield. Third, the model was extended to include a phenology scheme, a mass-balance equation relating mean leaf area index to accumulated GPP over the growth phase, and an independently observed response of leaf mass-per-area to CO_2. Sensitivity analyses and scenario runs with this extended model showed a positive but saturating (at ~600 ppm) response of yield to rising CO_2, consistent with experimental evidence. This positive effect was partially counteracted by a net negative response of yield to increasing temperature, caused by increasing photorespiration and an accelerated growth cycle.

Keywords

wheat; photosynthesis; crop yield; crop model; CO_2 fertilization; harvest index

1. Introduction

An adequate food supply is an essential basis for economic development and social stability in the context of increasing population and continuing anthropogenic climate change (Porter et al., 2014). As one of the world's four major crops (with maize, rice and soybean), wheat provides about a quarter of the world's cereal production (FAOSTAT, 2018) which, in turn, provides two-thirds of human caloric intake (Zhao et al., 2017). Wheat was introduced from the Near East and been cultivated in China since the late 6^{th} to early 5^{th} millennium BP (Betts et al., 2014). China is now both the largest producer and the largest consumer of wheat (Wang et al., 2009). Current wheat production in China exceeds 134 Mt of grain per year. This is more than 17% of the total global wheat production, and about 22% of the total cereal production of China (FAOSTAT, 2018). Thus, even a small fluctuation in China's wheat production could potentially impact not only the Chinese economy but also global food security.

The growth and harvestable yield of wheat are determined by environmental factors (Asseng et al., 2004) but also strongly influenced by management (He et al., 2015). Light, CO_2, temperature, water and nutrient availability define the basic conditions for the growth and development of the crop. Light and CO_2 directly affect photosynthesis (Gerbaud and Andre, 1980); temperature further influences growth and development processes including germination, anthesis and harvest (Asseng et al., 2011; He et al., 2015; Liu et al., 2018; Porter and Gawith, 1999; Tao et al., 2012); water and nutrient availability principally influence foliage cover (Nielsen and Halvorson, 1991; Pan et al., 2019) and therefore the absorption of light for photosynthesis. However, the basic conditions of wheat growth, especially temperature and CO_2 concentration, are changing. Temperatures in China have risen by 1.2˚C over the past few decades (Cao et al., 2017; Piao et al., 2010) and continued

66 warming is expected in the coming decades (Kirtman et al., 2014). Winter is warming faster than
67 summer (Piao et al., 2010; Wu et al., 2017) and this situation is potentially unfavorable to the
68 production of wheat (Brooking, 1996). On the other hand, atmospheric CO_2 already exceeds 400
69 ppm, more than 40% above its pre-industrial level and is expected to continue rising (Collins et al.,
70 2014). For C_3 crops (including wheat) the effect of rising CO_2 level on photosynthesis is positive
71 (Ainsworth and Rogers, 2007; Boylan, 2016; Sage et al., 1989), leading to higher photosynthetic
72 productivity and potentially also grain yield (Lawlor and Mitchell, 1991). Improved management
73 practices (e.g. fertilization, irrigation) and crop breeding have also contributed to increased wheat
74 yield (Qin et al., 2015; Yu et al., 2019) under current climate conditions, and this trend is expected to
75 continue.

76 The combined effects of changes in climate, CO_2 and management are highly uncertain
77 (Challinor and Wheeler, 2008) and numerical models are needed to project future trends in yield in
78 different regions, and thereby to facilitate adaptation in the food production system. Such models
79 should integrate knowledge from experiments and observations with theoretical understanding. Crop
80 models have been under development for at least 40 years, and there are now many models with the
81 technical capacity to simulate the growth and development of wheat (Blanc, 2017; Huang et al., 2016;
82 Palosuo et al., 2011). However, current crop models require detailed input information that is
83 challenging to collect over large scales and potentially subject to change in a dynamic environment.
84 Moreover, inter-model comparisons have revealed large differences between model predictions of
85 both current yields and future trends (Liu et al., 2019; Nelson et al., 2014; Ostberg et al., 2018). This
86 situation parallels that for natural vegetation models (Prentice et al., 2015), and suggests that current
87 crop models contain untested and potentially incorrect assumptions. Recently, however, progress has

been made in developing simpler vegetation models, based on theoretical principles but drawing extensively on empirical data to test each model component (Franklin et al., 2017; Wang et al., 2017). Here we adopt this novel approach to develop a prototype model to predict wheat growth and yield. Further, we investigate climate-change impacts on yield, and compare our results with other crop model predictions. This kind of analysis is required for the assessment of future food security (Liu et al., 2019; Rosenzweig et al., 2014).

The starting point for this model ('PC', for P crop: see Fig. 1) was the universal C_3 primary production model 'P' (Stocker et al., 2019; Wang et al., 2017). The P model is a theoretically derived and extensively tested light use efficiency (LUE) model that predicts gross primary production (GPP) as a function of climate, absorbed light and CO_2 on time steps of a week to a month. Further model development and testing of the PC model, presented here, used measurements of GPP, biomass and yield at research sites across the wheat-growing regions of China. In the first step of the analysis, the original P model was applied to predict GPP at two flux-tower sites situated in wheat crops in order to test its performance. In the second step, simple empirical equations were fitted to experimental data at several field research sites in order to relate accumulated GPP to above-ground biomass (AB), and AB to grain yield. These equations were combined with the P model to predict yields for different sites and years; these predictions were compared to observed yields. Uncertainties in predicted yields due to key parameters of the P model, and to the fitted coefficients of the empirical equations, were quantified. In the third step, the model was extended to represent the responses of yield to environmental change by the inclusion of (a) a general scheme to predict phenology, (b) a mass-balance equation quantifying leaf area index (LAI) consistent with a given GPP, and (c) an observed relationship between leaf mass-per-area (LMA) and the CO_2 concentration experienced

during crop growth. These extensions allowed modelled light absorption to be influenced by changes in growing-season length, and changes in modelled GPP to feed back to LAI. The extended model was applied at six field sites to project future wheat yields under different combinations of increasing CO_2 and temperature, in a sensitivity analysis for combinations of CO_2 and temperature increase, and in alternative scenario runs for future CO_2 and temperature change.

2. Material and methods

2.1 The P model

The P model is based on the standard biochemical model of C_3 photosynthesis (Farquhar et al., 1980), with additional formulations that allow photosynthetic capacities and stomatal behaviour to acclimate to environmental conditions on weekly to monthly time scales (Wang et al., 2017). Instantaneous photosynthetic rates according to the standard model are the lesser of the electron transport-limited rate (A_J) and the carboxylation-limited rate (A_C). A_C is proportional to the Rubisco capacity (V_{cmax}). In the P model, V_{cmax} is assumed to acclimate to growth conditions such that the two rates are co-limiting under average daytime conditions (Maire et al., 2012; Smith et al., 2019). A_J is proportional to the absorbed photosynthetic photon flux density (PPFD) at low PPFD, increasing with PPFD towards a light-saturated rate that is proportional to the electron-transport capacity (J_{max}). In the P model, acclimation of J_{max} parallels that of V_{cmax} and their ratio is set to maximize the benefit (A_J) minus the cost of maintaining J_{max}. Both A_C and A_J are functions of the leaf-internal CO_2 partial pressure (c_i), whose ratio (χ) to the ambient CO_2 partial pressure (c_a) is determined by stomatal responses to the relative rates of carbon gain and water loss. In the P model, χ is determined by the least-cost criterion (Prentice et al., 2014), which minimizes the combined costs of maintaining the

Rubisco and water transport capacities. The three constraints (on V_{cmax}, J_{max} and χ) lead to an expression for weekly to monthly GPP under field conditions that has the mathematical form of a LUE model. That is, accumulated GPP is proportional to absorbed PPFD:

$$\mathrm{GPP} = \Phi_0 I_{abs} m \sqrt{1 - \left(\frac{c^*}{m}\right)^{\frac{2}{3}}} \quad (1a)$$

where

$$m = \frac{c_i - \Gamma^*}{c_i + 2\Gamma^*} \quad (1b)$$

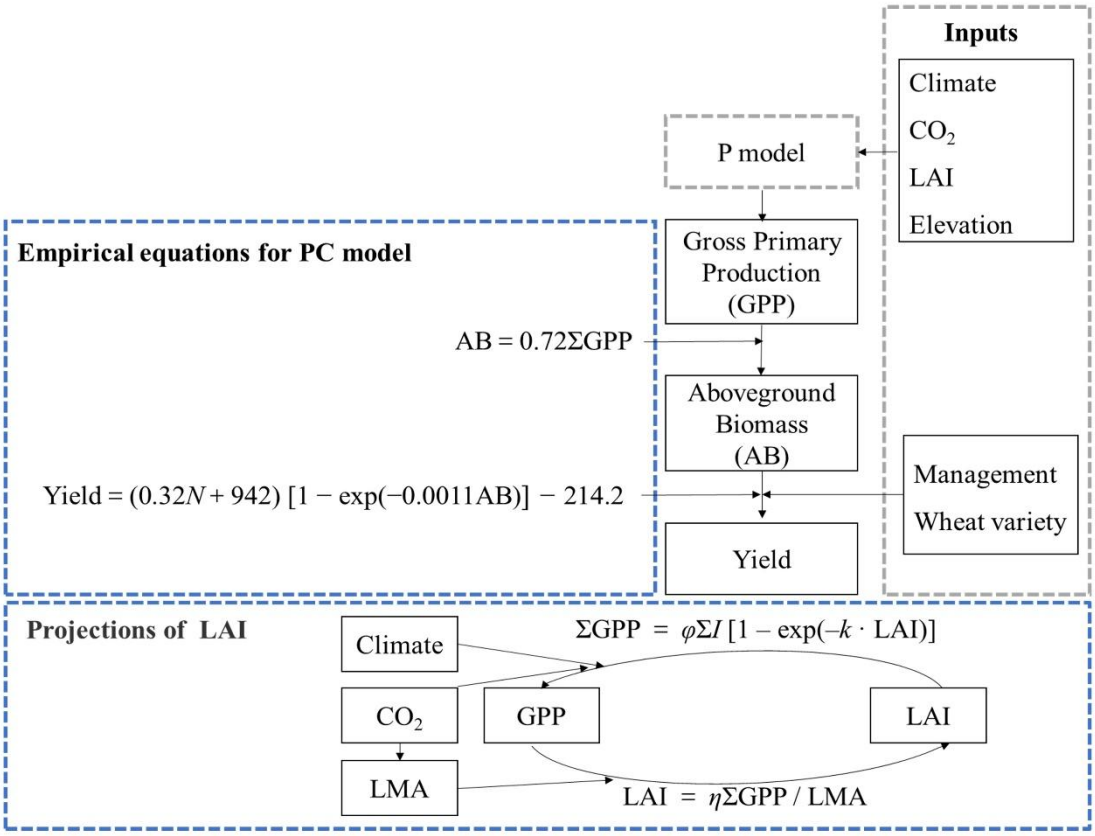
$$\frac{c_i}{c_a} = \frac{\Gamma^*}{c_a} + \frac{\xi\left(1 - \frac{\Gamma^*}{c_a}\right)}{\xi + \sqrt{D}} \quad (1c)$$

$$\xi = \sqrt{\frac{\beta\left(K + \Gamma^*\right)}{1.6\eta^*}} \quad (1d)$$

Here, Φ_0 is the intrinsic quantum yield (mol CO_2 mol^{-1}); I_{abs} is the PPFD intercepted and absorbed by the canopy (mol m^{-2} s^{-1}); c_a is the ambient atmospheric CO_2 partial pressure (Pa); Γ^* is the photorespiratory compensation point (Pa); η^* is the viscosity of water, relative to its value at 25 ˚C (dimensionless); D is the vapour pressure deficit (Pa); K is the effective Michaelis-Menten coefficient of Rubisco (Pa); and $c^* = 0.41$ and $\beta = 146$ are dimensionless constants, where c^* has been estimated from observed J_{max}:V_{max} ratios and β from observed stable carbon isotope ratios (Wang et al., 2017).

The P model thus calculates GPP as the product of I_{abs}, which is the product of incident PPFD and *f*APAR (the fraction of incident PPFD absorbed by foliage), and LUE. LUE is the product of Φ_0,

148 m and the square-root term in equation (1a). The parameters Γ^*, η^* and K depend on temperature
149 (Bernacchi et al., 2001; Wang et al., 2017) and Γ^* and K depend on atmospheric pressure (Farquhar
150 et al., 1980). The inputs to the model are air temperature (T), relative humidity (RH), incident PPFD,
151 fAPAR, elevation (to calculate atmospheric pressure) and c_a (the product of the current year's mole
152 fraction of CO_2 with atmospheric pressure). When driven by satellite-derived fAPAR data, the model
153 has been shown to reproduce monthly GPP well at eddy-covariance flux tower sites in natural
154 vegetation worldwide (Stocker et al., 2019; Wang et al., 2017) and geographic patterns, seasonal
155 cycles and interannual variability of GPP at flux tower sites located in different biomes, including
156 croplands (Balzarolo et al., 2018).



157

158 **Figure 1: The structure of the PC model.** ΣGPP: accumulated gross primary production over
159 growing season (g C m^{-2}); N: total application of nitrogen (kg N hm^{-2}). LAI: leaf area index
160 (dimensionless). φ: light use efficiency (%). $\sum I$: the sum of incident light over the growth phase (mol

161 photo m^{-2}). k: a dimensionless constant, (k = 0.5). η: the fraction of ΣGPP allocated to leaves
162 (dimensionless). LMA: leaf mass per area (g m^{-2}). Climate here comprises temperature, relative
163 humidity and incident photosynthetic photon flux density. Boxes outlined by grey dashed lines
164 represent the published model or known information, whereas the boxes with blue dashed lines show
165 components of the new model and its extension.

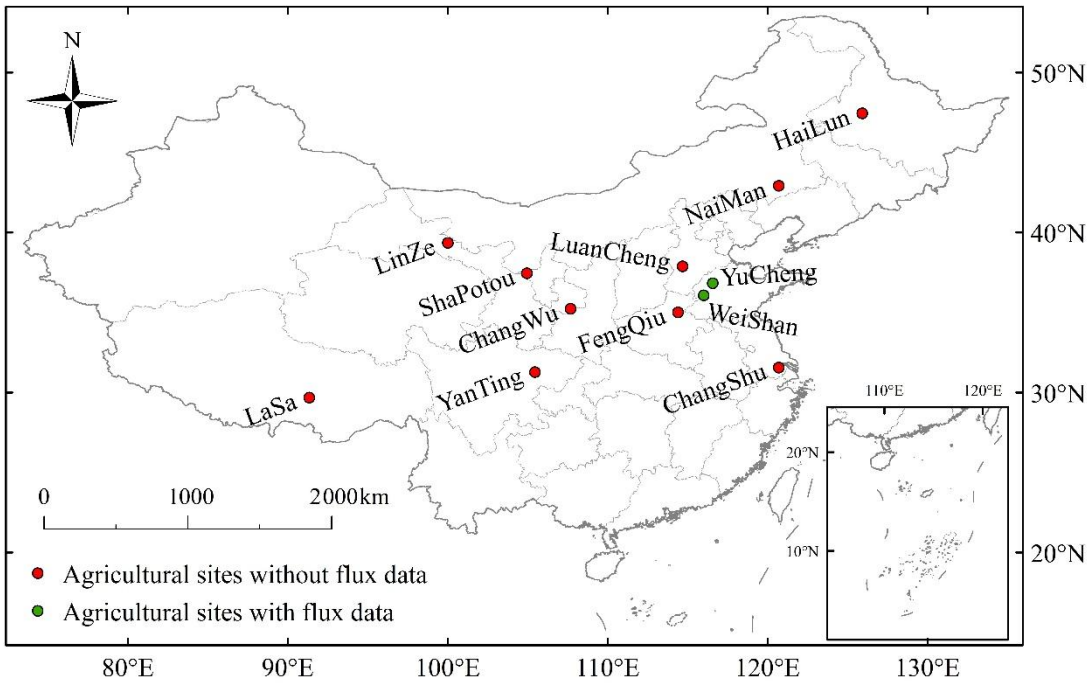
166 We ran the P model on a weekly time step. The model has already been shown to work well on a
167 ten-daily time step (Balzarolo et al., 2018). We applied the 'BRC' model set-up (Stocker et al., 2019),
168 which differs from the original published version (Wang et al., 2017) by incorporating an observed
169 temperature-dependence of Φ_0 (Bernacchi et al., 2003):

170 $$\Phi_0 = \frac{0.352 + 0.021T - 0.00034T^2}{8} \quad (2)$$

171 where T is temperature (˚C).

172 **2.2 Sites and field data**

173 Data from 12 agricultural sites in the main wheat-growing area of China (see Fig. 2) were used
174 for the second step of model development and testing (see Table 1 and Table 2). More than 90% of
175 wheat production occurs in the areas where these sites are located (Wang et al., 2009), so they are
176 representative of the environmental conditions for wheat production in China.



177

178 **Figure 2: Locations of sites providing experimental data.**

Site	Code	Longitude (E, °)	Latitude (N, °)	Elevation (m)	Mean annual air temperature (℃)	Mean annual precipitation (mm)	Potential heat units (PHU) (℃ day)
WeiShan	WS	116.83	36.23	20	13.3	532	_
YuCheng	YC	116.57	36.82	22	13.2	530	1676
ChangShu	CS	120.7	31.55	3.1	16.6	1321.2	1792
ChangWu	CW	107.68	35.23	1220	9.1	580	1549
LuanCheng	LC	114.68	37.88	50.1	12.2	536.8	1628
FengQiu	FQ	114.4	35	67	13.9	605	1797
YanTing	YG	105.45	31.27	420	17.3	836	1486
HaiLun	HL	125.92	47.45	240	4	550	_
LaSa	LA	91.33	29.67	3688	6	425	_
LinZe	LZ	100	39.35	1384	1.5	550	_
NaiMan	NM	120.7	42.92	358	5	425	_
ShaPotou	SP	104.95	37.45	1250	9.6	1250	_

179 **Table 1: Background information about the sites.** The value of PHU was set as the 10-year average of observations from 2005 to 2014.

Site code	Data span	PPFD (mol m^{-2} day^{-1})	T (℃)	RH (%)	CO_2 (ppm)	LAI	Elevation (m)	AB (g m^{-2})	Yield (g m^{-2})	Fertilization Irrigation	Wheat variety	Flux data (g C m^{-2} day^{-1})	Usage
WS	2006	√	√	√	√	√	√					√	a
YC	2004-2015	√	√	√	√	√	√	√	√	√	√	√*	a, b, c, d
CS	2004-2015	√	√	√	√	√	√	√	√	√	√		c, d
CW	2004-2015	√	√	√	√	√	√	√	√	√	√		c, d
FQ	2004-2015	√	√	√	√	√	√	√	√	√	√		c, d
LC	2004-2015	√	√	√	√	√	√	√	√	√	√		c, d
YG	2004-2015	√	√	√	√	√	√	√	√	√	√		c, d
HL	2005-2006		√				√	√	√	√	√		c
LS	2004-2015		√				√	√	√	√	√		c
LZ	2006		√				√	√	√	√	√		c
NM	2006		√				√	√	√	√	√		c
SP	2006		√				√	√	√	√	√		c

Table 2: Data details and use. √: data are available; a: sites used to test the P model; b: sites used to derive the GPP to AB relationship; c: sites used to derive the AB to yield relationship; d: were used to test the final model. * Two years (2004-2005) during the data span are available for the flux data at YuCheng.

183 Two flux tower sites (WeiShan, YuCheng; see Table 2 and Fig. 2) were used to test the GPP
184 predictions. One full year of observations from WeiShan (2006) and two years of observations from
185 YuCheng (2004, 2005) were available. The climate data (PPFD, *T* and RH) and canopy coverage
186 (here estimated from LAI by Beer's law), required as input to the P model, were obtained from on-
187 site measurements for WeiShan provided by the original authors (Lei and Yang, 2010) and
188 downloaded from the National Ecosystem Research Network of China (CNERN:
189 http://www.cnern.org.cn/) for YuCheng. CO_2 concentrations used are the annual global average
190 obtained from the United States National Oceanic & Atmospheric Administration (NOAA:
191 https://www.esrl.noaa.gov/gmd/ccgg/trends/).
192 There are no data on AB or grain yield from WeiShan, although this information is available for
193 YuCheng. We therefore used the two years of data from YuCheng to derive the relationship between
194 GPP and AB. We obtained experimental data from ten additional agricultural sites providing
195 information on AB and grain yield from CNERN. CNERN also provided data on climate (including
196 PPFD, *T* and RH), LAI, dates of the growing period, wheat varieties planted and management
197 practices (including irrigation and the supply of total nitrogen, phosphate and potassium) for all of
198 these sites. However, the records cover different periods (see Table 2): some sites only have data for
199 one year (LZ, SP, NM); one site has data for two years (HL); the remaining sites have records
200 spanning multiple years (CS, CW, FQ, LC, YG, LS). We used all the available data from these ten
201 sites and YuCheng together (584 data points) to estimate the allocation relationship between AB and
202 grain yield. We used data from six sites (CS, CW, FQ, LC, YG, YC) with records for more than two
203 years to test the final model. We could not use the Lasa site for testing because there are only
204 biomass and yield data, not LAI or climate data, from this site.
205 Climate data were pre-processed on a weekly time step, with PPFD summed and *T* and RH
206 averaged. Then vapour pressure deficit (*D*, kpa) was calculated according to the following equation
207 (Campbell and Norman, 2012):

$$D = 610.8 \exp\left(\frac{17.27T}{T + 237.3}\right)\frac{(100 - \mathrm{RH})}{100} \quad (3)$$

The fraction of absorbed photosynthetically active radiation was estimated by Beer's law (Monsi, 1953) from LAI:

$$f\mathrm{APAR} = 1 - \exp(-k \cdot \mathrm{LAI}) \quad (4)$$

where k is a dimensionless constant, assigned a generic value of 0.5. LAI was measured several times over the growing season, but the times of measurement varied from year to year and site to site. The LAI data used as input to test the P model (WS 2006, YC 2004-2005) are based on eight to ten observations at each site over the growing season. We interpolated the data to weekly mean values using a polynomial regression of LAI against time. Measurements of LAI at the sites used to test the crop model (PC model) were made more sporadically (less than five observations per growing season) and therefore could not be interpolated using polynomial regression. At these test sites, we used two methods to obtain the LAI input. First, we used the LAI scheme derived from LPJmL4 and integrated the mass-balance equation (see section 2.5.1) to calculate LAI. Second, we derived LAI values from the MODIS LAI product (MCD15A3H: 4-day time-step and 500m resolution, https://modis.gsfc.nasa.gov/). Since MODIS severely underestimates the observed LAI at the six test sites, we calculated the relationship between observed and MODIS LAI by linear regression and used the slope of this regression to rescale the MODIS LAI data and derive weekly mean LAI. We used calculated LAI and corrected MODIS LAI as alternatives to drive the PC model, and quantified the impacts of the choice of LAI inputs on model predictions.

2.3 Derivation of allocation relationships

We hypothesized that a fixed proportion (ε) of accumulated GPP during the growing season would be allocated to above-ground biomass (AB). We calculated GPP accumulation (ΣGPP) from

230 the beginning of the growing season to the day when AB was measured, then derived an empirical
231 relationship between AB and ΣGPP by linear regression. We used the slope of this linear regression
232 as an estimate of ε.

233 We hypothesized that grain yield should increase, monotonically but not necessarily linearly,
234 with AB, and that this relationship might be influenced by wheat varieties and management practices.
235 Non-linear regression was used to derive an empirical relationship between grain yield and AB,
236 taking account of the effect of nitrogen supply and wheat variety on the relationship, meanwhile
237 testing the effects of irrigation and the application of phosphate and potash. Non-linear regression
238 was performed using a mixed-effects model in the **nlm** package of R. The form of the fitted equation
239 is as follows:

240 $$\text{Yield} = \left(a \cdot N + b\right)\left[1 - \exp\left(c \cdot \text{AB}\right)\right] + d \quad (5)$$

241 where N is the total application of nitrogen (kg N hm^{-2}), and a, b, c and d are parameters to be fitted.
242 Wheat variety was considered as a random effect added to parameter b, thus allowing maximum
243 yields to differ by variety. The potential effects of other factors (irrigation, precipitation, mean
244 temperature during the growth season) were tested by examining the residuals from this regression.
245 To check the goodness of fit of the non-linear regression, linear regression was also performed both
246 using all the data together, and for each variety separately. The root mean squared error (RMSE) and
247 Akaike Information Criterion (AIC) were calculated as an indicator of the goodness of fit of each
248 model.

249 **2.4 Model evaluation**

250 We tested the performance of the PC model by comparing interannual variations in predicted
251 and observed AB and grain yield over multiple years at the six test sites (CS, CW, FQ, LC, YC, YG),
252 using meteorological observations from each site to drive the model. The simulated accumulated

GPP during the growing season was allocated to AB using the fixed ratio (ε) obtained by linear regression, and AB at harvest was allocated to grain yield using the non-linear relationship described above (Equation 5). The growing season was defined as the period when mean daily temperature was above 0°C. Interannual variation in yield provides an independent test of the model, as no information on interannual variability was used in the derivation of empirical relationships used in the model. Additional metrics – mean absolute error (MAE), Nash-Sutcliffe Efficiency (NSE) and RMSE – were calculated to evaluate model performance.

There are two sources of uncertainty in the model predictions: the input data (climate and LAI) and the model (model parameters and model errors). For the formal uncertainty analysis, we assumed the input data were reliable and focused on uncertainty due to model parameters. Three aspects of uncertainty were considered, as follows:

- Equation 1: The two most uncertain quantities (β, c^*) in the P model (Prentice and Thomas, 2018). β is the ratio of the unit costs for the maintenance of carboxylation and transpiration capacities, evaluated at 25˚C. It determines the value of the ratio of leaf-internal to ambient CO_2 (an index of stomatal behaviour) under standard conditions. It was estimated, based on global leaf $\delta^{13}C$ data, as $\beta = 146 \pm 2.7$ (Wang et al., 2017). c^* is the unit cost of maintaining electron-transport capacity and determines the extent to which optimum carboxylation capacity is lowered because of the cost of maintaining an equivalent capacity for electron transport. It was estimated, based on experimental measurements under a variety of growth conditions, as $c^* = 0.41 \pm 0.112$ (Wang et al., 2017).

- The proportionality constant (ε) between biomass and accumulated GPP. This describes how much accumulated GPP is allocated to above-ground biomass over growing season. It was estimated based on the observation of accumulated GPP and above-ground biomass at YuCheng

site in 2004-2005, from the linear regression of above-ground biomass against accumulated GPP. The uncertainty of ε was assessed from the standard error of the slope. The value used was ε = 0.72 ± 0.032.

- Equation 5: The four main parameters (a, b, c, d) from the formula relating yield to biomass. These describe the effects of above-ground biomass and nitrogen on yield, and were estimated by non-linear regression based on observations of above-ground biomass and yield at several agricultural sites (see Table 2). Their uncertainties were assessed from the standard error of the estimated values obtained from a mixed-effects model. The value used were a = 0.32 ± 0.11, b = 942 ± 50.49, c = −0.0011 ± 0.00013 and d = −214.2 ± 59.26.

We considered each of the source of uncertainty in the individual equations in the PC model independently, and combined these uncertainties using the standard error propagation formula (JCGM, 2008; Prentice and Thomas, 2018):

$$u^2(y) = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + 2u_c(y) \quad (6a)$$

$$u_c(y) = \sum_{i=1}^{n-1} \sum_{i+1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_{i+1}} u(x_i) u(x_{i+1}) r(x_i, x_{i+1}) \quad (6b)$$

where $u(y)$ is the standard uncertainty (of GPP or AB or yield); n is the number of parameters; $\partial f / \partial x_i$ is the sensitivity to parameter x_i (obtained by differentiating the individual equations); $u(x_i)$ is the standard uncertainty of x_i; $u_c(y)$ is the uncertainty derived from the correlation among parameters, and $r(x_i, x_{i+1})$ is the degree of correlation between parameter x_i and parameter x_{i+1}. If parameter x_i and parameter x_{i+1} are independent, $r(x_i, x_{i+1}) = 0$.

2.5 Model extension

2.5.1 Phenology scheme

The phenology scheme for wheat was adopted from the LPJmL4 model (Bondeau et al., 2007; Schaphoff et al., 2018). Sowing and maturity dates were obtained from the datasets provided in the global gridded crop model inter-comparison project (Elliott et al., 2015). A phenological scalar (f_{PHU}) ranging from 0 at sowing to 1 at harvest was computed:

$$f_{\mathrm{PHU}} = \sum_{1}^{n} \frac{T_{\mathrm{m}} - T_{\mathrm{b}}}{\mathrm{PHU}} \tag{7}$$

where n is the number of days from sowing. PHU is the number of phenological heat units, and describes the total heat requirement over the growing season. Its value was set as the 10-year average of observations from 2005 to 2014 for each site. T_m is the daily mean air temperature (˚C) and T_b is the base temperature (here 0˚C). LAI development was determined using a sigmoid curve during the growth phase and a quadratic curve during the senescence phase. (In the absence of water and nutrient stresses, LAI is assumed to follow this optimal curve.) During the growth phase:

$$f_{\mathrm{LAImax}} = \frac{f_{\mathrm{PHU}}}{f_{\mathrm{PHU}} + \exp\left(l_1 - l_2 \cdot f_{\mathrm{PHU}}\right)} \tag{8}$$

where LAI is a time-dependent fraction (f_{LAImax}) of maximum LAI ($\mathrm{LAI}_{\mathrm{max}}$), and l_1 and l_2 are the first and second inflection points. During the senescence phase:

$$f_{\mathrm{LAImax}} = \frac{\left(1 - f_{\mathrm{PHU}}\right)^2}{\left(1 - f_{\mathrm{PHUsen}}\right)^2}\left(1 - f_{\mathrm{LAImax\text{-}harvest}}\right) + f_{\mathrm{LAImax\text{-}harvest}} \tag{9}$$

where f_{PHUsen} is the fraction of PHU when senescence begins, and $f_{\mathrm{LAImax\text{-}harvest}}$ is the fraction of maximum LAI at harvest (here fixed to zero). The f_{PHU} values corresponding to the l_1 and l_2

314 inflection points were derived from Figure 4 in Bondeau et al. (2007). The values of l_1 and l_2 were
315 calculated for these f_{PHU} values, using the method of Neitsch et al. (2011), as 0.89 and 10 respectively.

316 2.5.2 Dynamic leaf area index

317 Prognostic calculation of LAI was enabled by fitting a linear relationship between leaf biomass
318 and AB based on data from the field sites, then solving for LAI in the mass-balance equation:

319 $$\frac{\mathrm{LMA}}{\eta}\cdot\mathrm{LAI}=\varphi\sum I\left[1-\exp\left(1-k\cdot\mathrm{LAI}\right)\right] \quad (10)$$

320 where η is the fraction of ΣGPP allocated to leaves, LMA is the leaf mass-per-area (g m^{-2}), φ is the
321 modelled LUE (the ratio of modelled GPP, following equation (1), to absorbed PPFD), and ΣI is the
322 accumulated incident PPFD (mol photon m^{-2}). k is a dimensionless constant (k = 0.5). Equation (10)
323 indicates that the LAI demand (left hand side) must equal its supply (right hand side). The LAI
324 demand represents the allocation of accumulated GPP to the canopy to support a given level of LAI.
325 The LAI supply represents the carbon accumulation supported by a certain level of LAI. η was
326 estimated from the observed data on leaf biomass and ΣGPP from the YuCheng site in 2004 and
327 2005, LMA was set at 35.7 g m^{-2} corresponding to the mean observed value at YuCheng, and
328 allowed to increase linearly with CO_2 using the observed slope (0.05 g m^{-2} per ppm) from
329 Thilakarathne et al. (2013).

330 **2.6 Model application**

331 2.6.1 Sensitivity analysis

332 Using 2005 as a baseline (baseline temperature is the weekly mean temperature over the
333 growing season and baseline CO_2 is 380 ppm), we ran simulations with the extended model,
334 including prognostic phenology and dynamic LAI and LMA, with temperature increasing by 0.05°
335 increments up to 5° above the baseline temperature and CO_2 concentration increasing by increments

of 5 ppm up to 500 ppm above the baseline CO_2 concentration. These changes were superimposed on the weekly mean temperatures and on the annual CO_2 concentration. All other inputs (radiation, relative humidity, management practices and wheat variety) were fixed at their 2005 values.

2.6.2 Future projections and comparisons with other models

We used the model to examine the consequences of potential future climate changes on wheat yields, following the protocol used by the Inter-Sectoral Impact Model Intercomparison Project-2b (ISIMIP2b: https://www.isimip.org/protocol/#isimip2b/). Climate data, including daily mean temperature, photosynthetically active radiation (assumed to be half of downward shortwave radiation) and relative humidity from the MIROC5 simulations, and CO_2 concentrations, for two scenarios (RCP2.6 and RCP6.0) were used to run the PC model at six test sites (CS, CW, FQ, LC, YC, YG) from 2006 to 2099. Management practices and wheat varieties were unchanged from 2005.

The LPJmL (Bondeau et al., 2007; Muller and Robertson, 2014; Schaphoff et al., 2018), GEPIC (Liu et al., 2007) and PEPIC (Liu et al., 2016) crop models have been used to simulate future wheat yields in ISIMIP2b. We compared our future projections of yield with results from these three models. In ISIMIP2b, these models ran simulations with full irrigation and no irrigation. In order to eliminate the effect of variable water supply, we compared our results with those from the full-irrigation run. We extracted simulated wheat yields at our six test sites from the results of each model for the RCP2.6 and RCP6.0 scenarios. Further information about these three models, including input data, leaf area, phenological development, yield formulation and stresses considered, is given in Table S1.

3. Results

3.1 Modelled versus observed GPP

Predicted weekly GPP values were consistent with the observations from the flux towers, both in their magnitudes (Fig. 3a) and their patterns over the growing season (Fig. 3b). Observed and

predicted GPP were highly correlated (R^2 = 0.81, RMSE = 10.1 g C m^{-2} $week^{-1}$) and the slope of the relationship was close to 1:1 (slope = 1.07 ± 0.08) with a non-significant offset (intercept = 0.79 ± 4.67 g C m^{-2} $week^{-1}$).

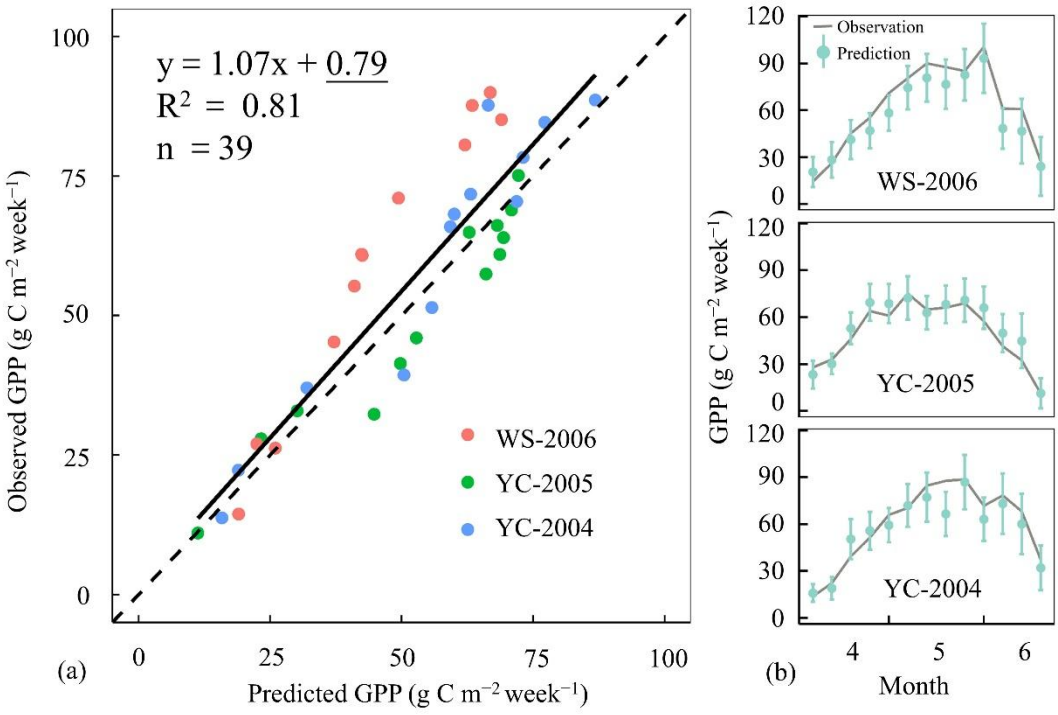
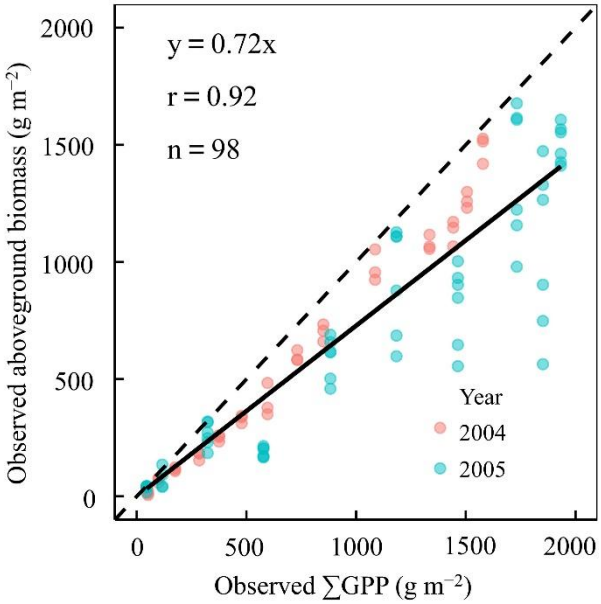


Figure 3: Comparison of predicted and observed gross primary production at two sites. (a) Scatterplot. The thick black line is the linear regression. Underscore means non-significant. **(b)** GPP during the growing season (weekly sums). WS-2006, YC-2005 and YC-2004, represents WeiShan in 2006, YuCheng in 2005 and YuCheng in 2004, respectively.

3.2 The relationship between AB and GPP

A strong linear relationship (Fig. 4) was shown between AB and accumulated GPP (r = 0.92) with an estimated 72% (slope, ε = 0.72) of accumulated GPP allocated to AB. The intercept was statistically significant, but small enough to be neglected.



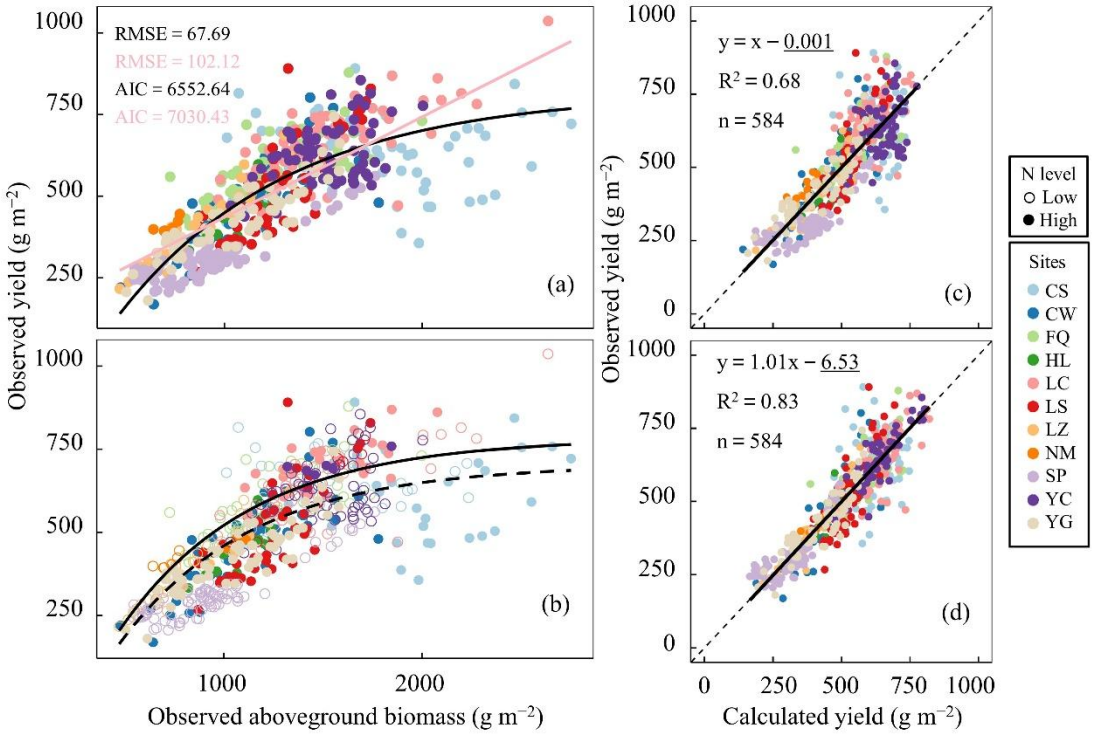
373

374 **Figure 4: The proportion of accumulated gross primary production (GPP) allocated to**
375 **above-ground biomass.** The data are observations during the wheat growth season at YuCheng from
376 2004 to 2005. All values were accumulated from green-up to measurement time.

377 **3.3 The relationship between yield and AB**

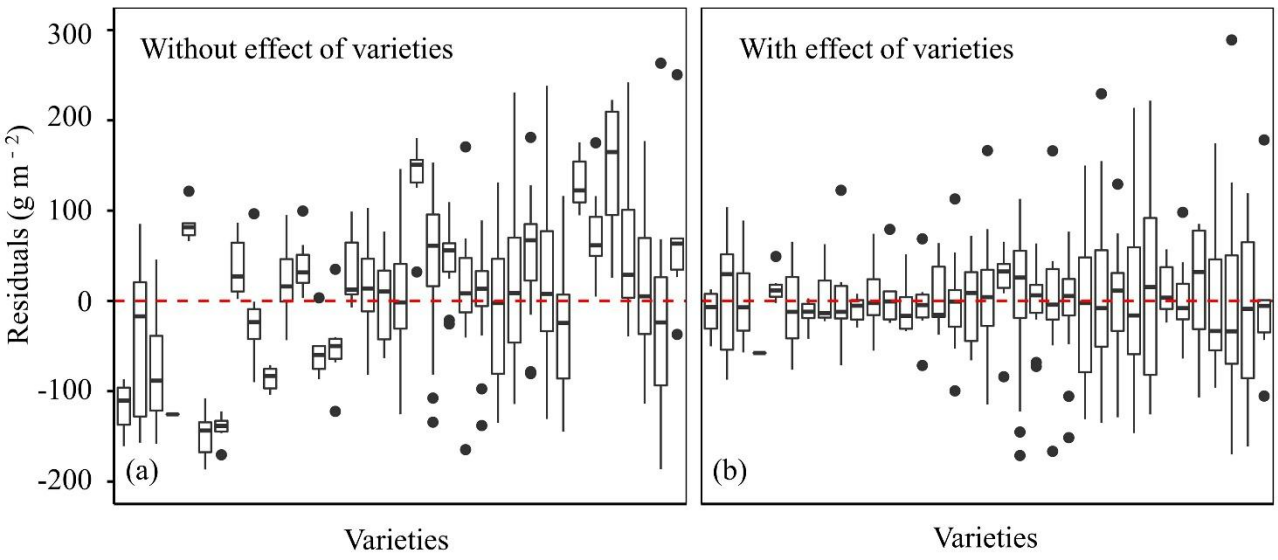
378 Yield was shown to be a saturating function of AB (Fig. 5a). Nitrogen addition affected the
379 overall level of allocation (Fig. 5b), with higher nitrogen supply causing high allocation to AB. The
380 relationship was affected by wheat variety, and a saturating relationship can also be shown within
381 each of the varieties that covers a large range of AB (range > 1800 g m^{-2}), with a substantially
382 smaller RMSE and AIC compared to linear fits (Fig. S1). Moreover, the slopes of linear regressions
383 fitted to each variety separately decline with the mean value of AB for the variety (Fig. 7 and Fig.
384 S2). In other words, at the high end of AB values, the increment in yield diminishes with the
385 increment in AB. These results indicate that the non-linear, saturating relationship of yield to AB
386 applies generally, both within and across varieties.

387



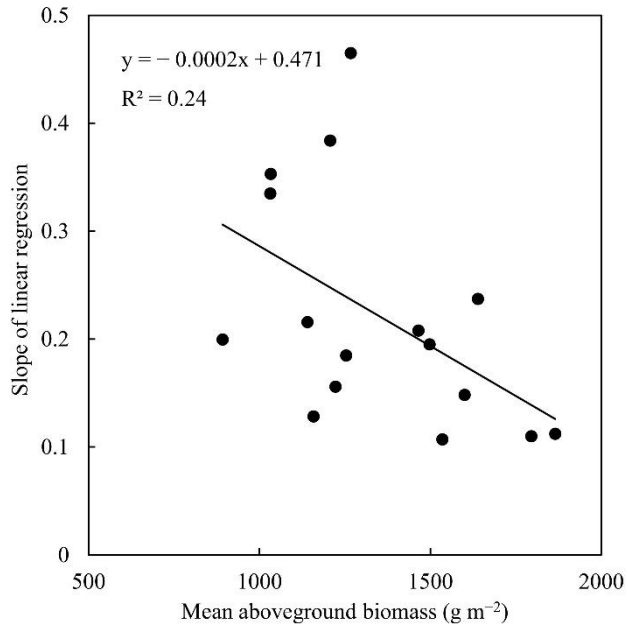
388

389 **Figure 5: Results of the mixed-effects model. (a)** Yield versus above-ground biomass (AB).
390 Pink line is linear regression and black line is non-linear regression. **(b)** Yield trend with AB,
391 including the effect of nitrogen addition. The solid line is the fitted curve of yield with AB at high
392 nitrogen level (pure nitrogen added = 300 kg hm^{-2}) and the dotted line is at low nitrogen level (pure
393 nitrogen added = 100 kg hm^{-2}). The open and closed circles represent the observations with total
394 application of pure nitrogen at levels above and below 200 kg hm^{-2}, respectively. **(c)** Scatterplot
395 including AB and nitrogen as predictors. **(d)** Scatterplot including AB, nitrogen and variety as
396 predictors. Underscore means non-significant.



397

398 **Figure 6: The relationship between residuals and wheat varieties. (a)** Without the random
399 effect of variety; **(b)** with this effect. Each box represents a wheat variety. The red line is zero and the
400 black dots are outliers.



401

402 **Figure 7: The fitted slope and mean value of above-ground biomass based on the linear**
403 **regression of yield against above-ground biomass within each variety.** (See Fig. S2 for the
404 separate linear regressions.)

405 The comparison between simulated and observed yields improved when variety was taken into
406 account (Fig. 5c compared to Fig. 5d). Residuals of the non-linear regression were reduced (Fig. 6),
407 and the correlation between predicted and observed yield improved (R^2 increased from 0.68 to 0.83).
408 Irrigation, mean temperature over the growing season, and the supply of phosphate and potassium
409 had no significant effects on the relationship between AB and yield ($P > 0.05$) on yield (Fig. S3),
410 indicating that the first-order effects of these factors are already subsumed in AB.

411 **3.4 Model evaluation**

412 3.4.1 Prediction of AB and yield variations

413 We used two alternative LAI inputs (calculated LAI, and corrected MODIS LAI) to drive the PC
414 model. The model captured the pattern of interannual variation in AB (see Fig. 8) and, although there
415 were some anomalous years, the predicted AB was generally within the range of the observations.
416 The simulated and observed yields matched reasonably well and interannual variations in grain yield
417 were captured (Fig. 9), with observations almost always within the uncertainty range of the
418 predictions. The choice of LAI input however affected model performance. Calculated LAI produced
419 better results than corrected MODIS LAI for the prediction of AB and yield (Figs 8–9), with
420 substantially larger values of NSE (0.75 and 0.77) and smaller values of MAE (113.50 g m^{-2} and
421 42.80 g m^{-2}) and RMSE (153.79g m^{-2} and 57.08 g m^{-2}). The corresponding values obtained with
422 corrected MODIS LAI were: NSE (–0.20 and 0.27), MAE (247.30 and 73.93 g m^{-2}) and RMSE
423 (313.66 and 93.34 g m^{-2}).

424 The most likely explanation for the low model efficiency of predictions driven by corrected
425 MODIS LAI is the uncertainty of MODIS LAI itself. Many previous studies have pointed to a large
426 underestimation of LAI by the MODIS product, compared to LAI values measured *in situ* (Fang et
427 al., 2019; Fu et al., 2016; Sharp et al., 2018), with multiple causes (Yang et al., 2007). In addition,

remote sensing of LAI in croplands suffers from specific problems due to their fragmented distribution and close association with settlements. Our correction procedure could not fully eliminate these problems.

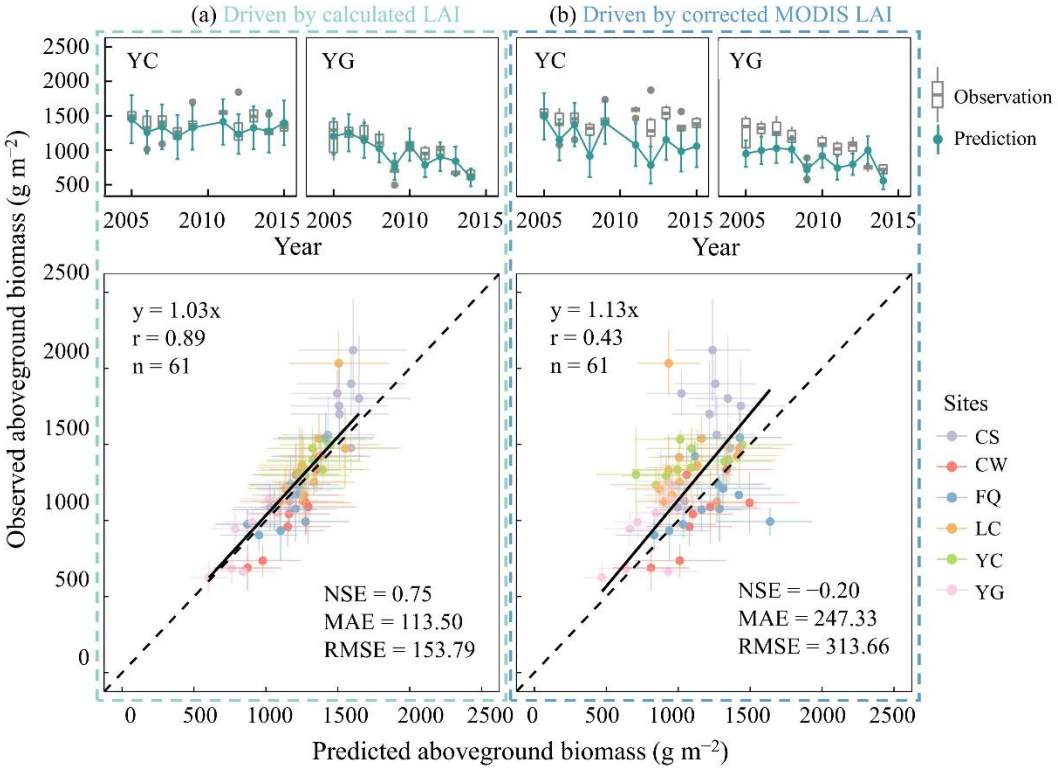
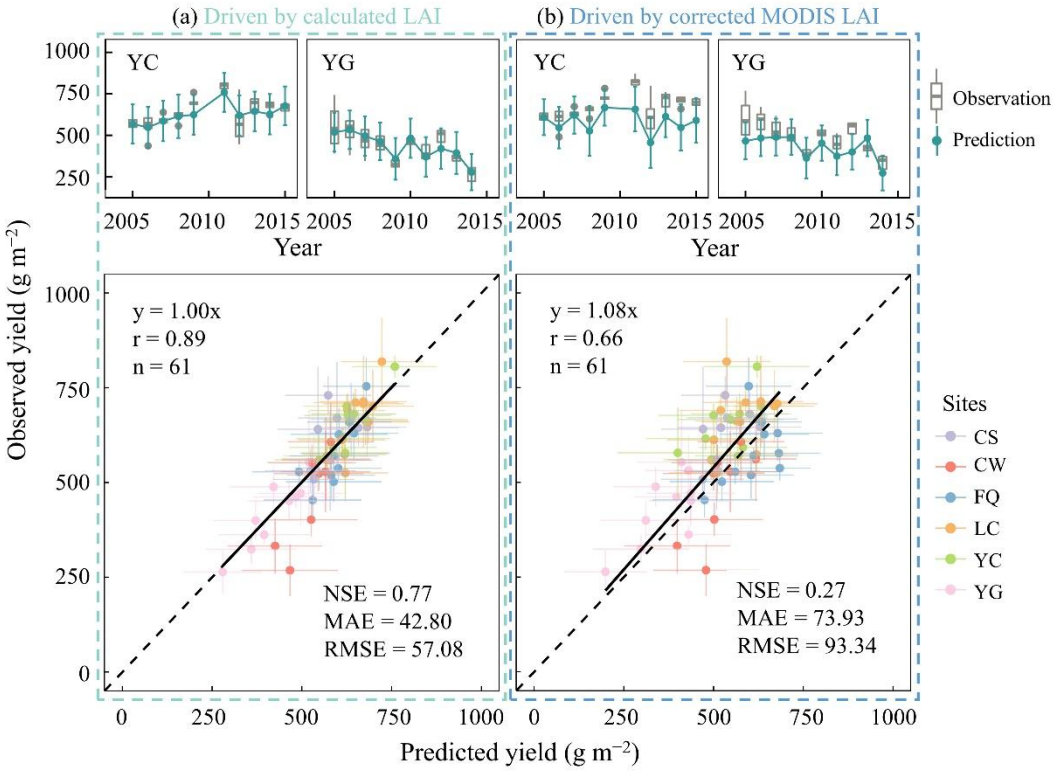


Figure 8: Comparisons of observed and modelled above-ground biomass. (a) Using LAI calculated by the scheme from LPJmL4 to drive the PC model. **(b)** Using corrected MODIS LAI to drive the PC model. NSE, Nash-Sutcliffe efficiency; MAE, mean absolute error; RMSE, root mean squared error.



436

437 **Figure 9: Comparisons of observed and modelled yield. (a)** Using LAI calculated by the
438 scheme from LPJmL4 to drive the PC model. **(b)** Using corrected MODIS LAI to drive the PC
439 model. NSE, Nash-Sutcliffe efficiency; MAE, mean absolute error; RMSE, root mean squared error.

440 3.4.2 Uncertainty analysis

441 Uncertainty in model predictions could originate either in the input data (climate, LAI) or in the
442 model. In this uncertainty analysis, we assumed that the input data were reliable, and used YuCheng
443 2005 as a case study to analyse the uncertainties due to the following model parameters:

444 • The two most uncertain quantities (β, c^*) in the P model (Equation 1),

445 • the proportionality constant (ε) between biomass and accumulated GPP, and

446 • the four parameters (a, b, c, d) from the formula relating yield to biomass (Equation 5).

447 The calculated uncertainties with respect to different model parameters for predicted GPP and
448 yield are shown in Fig. 10a. The total uncertainty of predicted GPP was ~24%, almost entirely

associated with the parameter c^*, which is an important control on the magnitude of simulated GPP. This parameter also contributed substantially (~14.6%) to the uncertainty in simulated yield. Additional substantial (~0.5%) contributions to the total uncertainty of modelled yield come from the four main parameters (a, b, c, d) from the formula relating yield to AB, which control the yield attained for a given biomass. The total uncertainty of predicted yield was ~15.3%. The correlations among parameters are summarized (Fig. 10b), together with the uncertainties and sensitivities of predicted yield for each model parameter (Fig. 10c). The parameter c^*, with the third largest uncertainty (~27%) and the second largest sensitivity (~5%), is the largest source of uncertainty in modelled yield (more than 90% of the total uncertainty). Three parameters (c^*, a, d) have large (>25%) uncertainties. A ± 10% change in the parameter b has the greatest effect on predicted yield, but its uncertainty is relatively small. Two parameters (β, c^*) in the P model, and the constant (ε) relating AB and accumulated GPP, were estimated independently and are therefore uncorrelated. The correlations between parameters a, b, c, d are shown in Fig. 10b. Parameter b has a strong positive correlation (r = 0.895) with c, and a strong negative correlation (r = −0.686) with d.

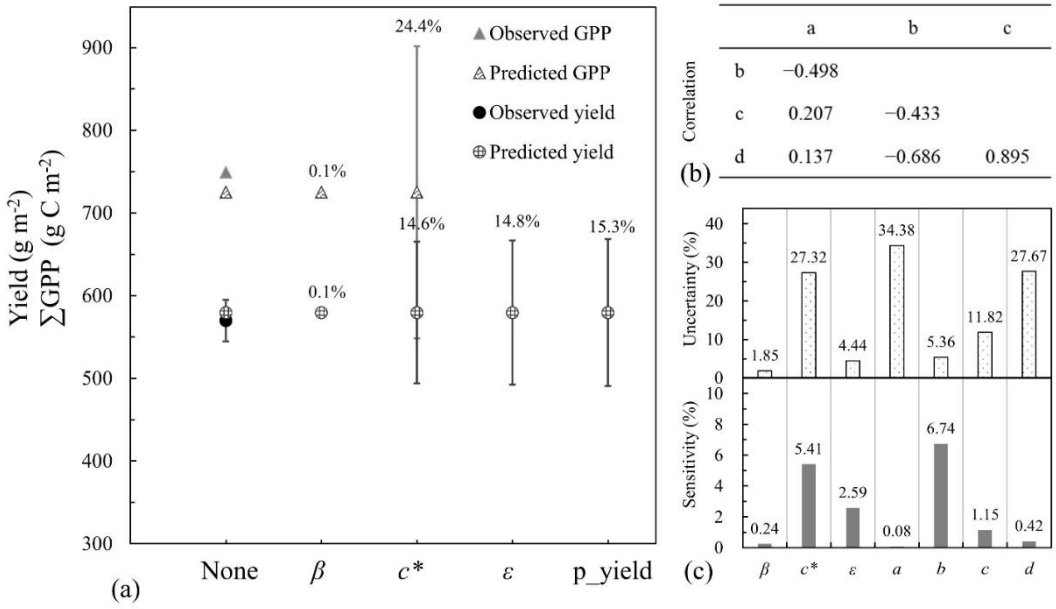


Figure 10: The impact of parameter uncertainty on the prediction of GPP accumulation and yield. YuCheng 2005 was selected as a case study. **(a)** The cumulative contribution of different parameters to uncertainty in modelled GPP and yield. Triangles represent GPP and circles represent yield. The *x*-axis represents the progressive inclusion of successive parameters, indicated by their mathematical symbols. p_yield represents the total effect of the four parameters from eq. (5). **(b)** The table of correlations among the four parameters (*a*, *b*, *c*, *d*) from eq. (5) relating yield to above-ground biomass **(c)** The uncertainty of each parameter (top); and the sensitivity of each parameter (bottom), defined as the change in modelled yield due to a ±10% change in the parameter value.

3.5 Model extension and application

3.5.1 Testing the phenology scheme

The phenology scheme was shown to reproduce seasonal patterns of LAI in the current climate (Fig. 11) reasonably well ($R^2 = 0.84$).

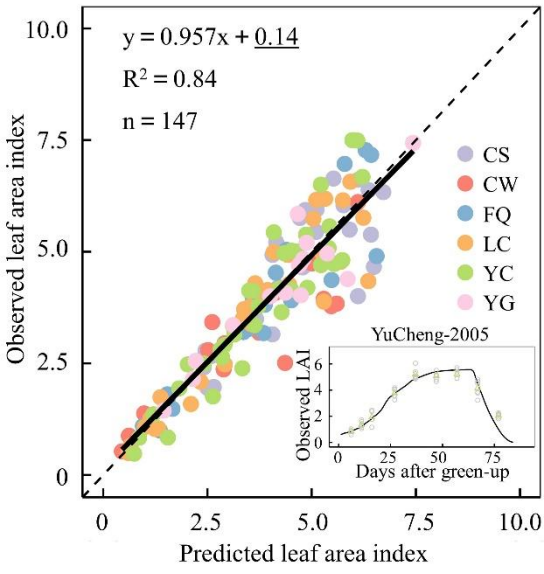


Figure 11: Observed versus predicted leaf area index (LAI) with the LPJmL4 phenology scheme. The inset shows the observed (circles, green is mean value) and predicted (line) seasonal time course of LAI at the YuCheng site in 2005. Underscore means non-significant.

3.5.2 The relationship between leaf biomass and accumulated GPP

A strong linear relationship (r = 0.94) was found between leaf biomass and accumulated GPP, allowing us to estimate η = 0.32 (Fig. 12) and thereby solve equation (10) for mean LAI over the growth period.

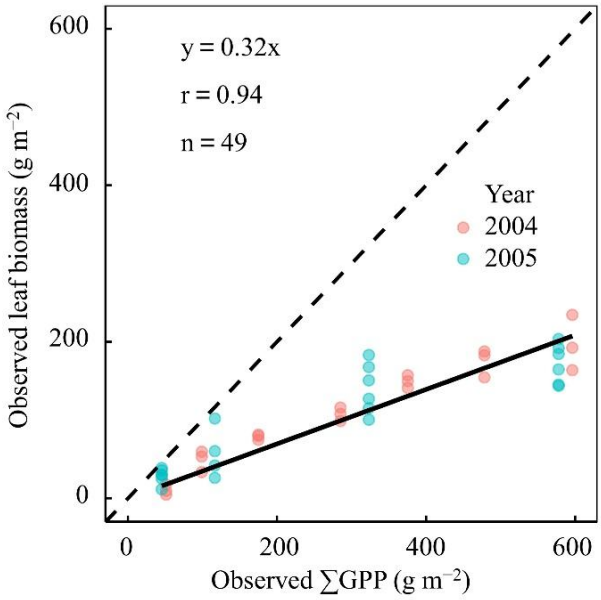


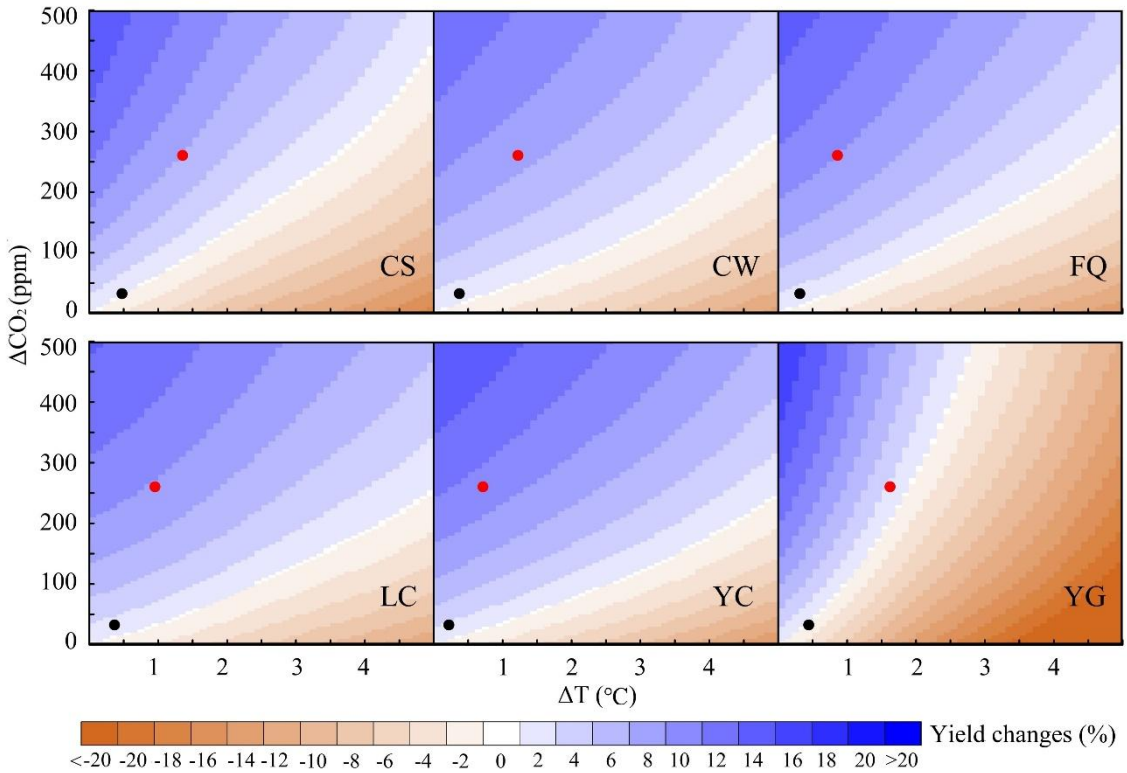
Figure 12: The proportion of gross primary production (GPP) allocated to leaves. The data are observations during the wheat growth phase in 2004 and 2005 at YuCheng. All values were accumulated from green-up to measurement time.

Projections of changing LAI as a function of CO_2 concentration are shown in Figs 14b and S4. Modelled LAI increases in response to increasing CO_2, but when the effect of increasing LAI on LMA is considered, the increase is much smaller and reaches a maximum at around 600 ppm. This behaviour is consistent with the maximum yield enhancement indicated by raised CO_2 experiments, as summarized in the meta-analysis by Broberg et al. (2019).

493 3.5.3 Sensitivity analyses

494 Modelled grain yields increase with rising CO_2 concentration, and decrease with increasing
495 temperature, when other variables are kept fixed (Fig. 13). Higher temperature shortens the growing
496 season (Fig. 14d) leading to a reduction in total absorbed light. In addition, the response of modelled
497 LUE to rising temperature follows a unimodal curve (see Fig. 14c) such that increasing temperature
498 above the optimum reduces GPP (Long, 1991). Lower GPP means lower yield. On the other hand,
499 increasing CO_2 monotonically increases LUE, GPP and yield; and increasing GPP leads to increasing
500 LAI, amplifying this effect. But the net effect of CO_2 is limited by an increase in LMA.

501 For these scenarios and sites, the modelled positive effect of rising CO_2 concentration on yield
502 was greater than the negative effect of increasing temperature. However, the modelled reductions in
503 yield caused by rising temperature differed among the sites (Fig. 13). Modelled wheat yields in
504 warmer regions today, like YG and CS, are more sensitive to warming than cooler regions such as
505 LC.



506

Figure 13: The response of predicted yield to rising CO_2 and increasing temperature at six sites. Dots show the increases of temperature and CO_2 concentration in the last decade (2090-2099) compared with the first decade (2006-2015) under two scenarios (RCP2.6 black dot; RCP6.0, red). The temperature is the mean value over the growing season.

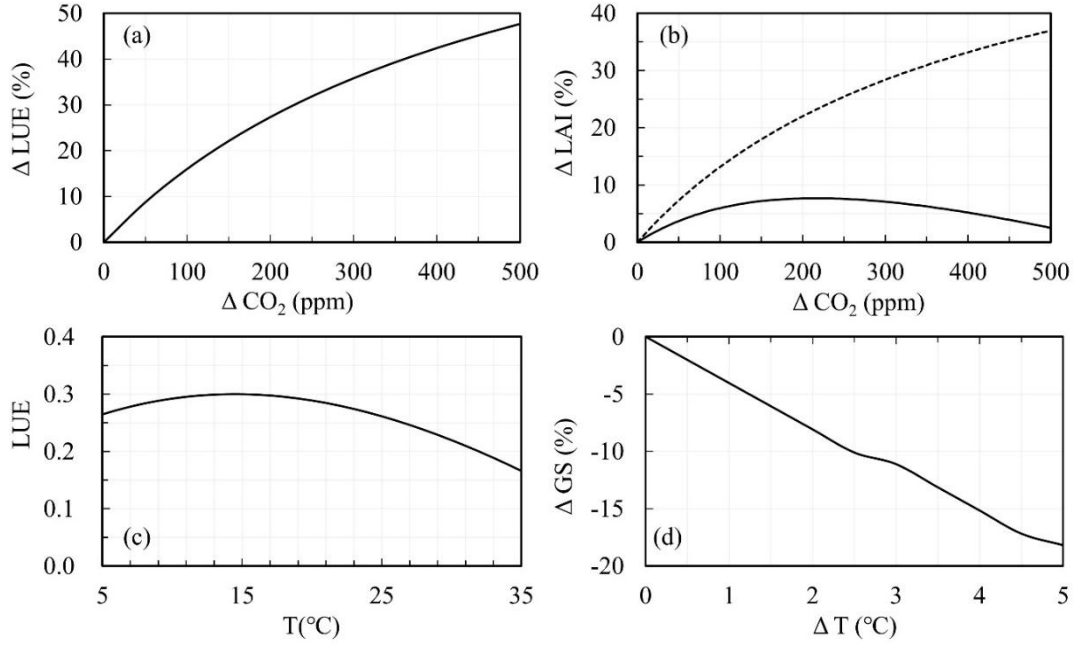


Figure 14: The response of light use efficiency (LUE), leaf area index (LAI) and growing season length (GS) to CO_2 concentration and temperature (T). Using the climate and CO_2 measurements from YuCheng 2005 as a baseline condition. **(a)** The response of LUE change to CO_2 increment. **(b)** The response of LAI change to CO_2 increment. Solid line includes the effect of CO_2 on LMA; dash line excludes this effect. **(c)** The response of LUE to temperature. **(d)** The response of changes in growing season length to temperature increment.

3.5.4 Comparison with future scenario runs from other crop models

We compared future scenarios with the PC model to ISIMIP2b model runs performed with the same scenarios using complex crop models. PC and LPJmL simulated contemporary yields reasonably well across all the sites, but the PEPIC and GEPIC models showed unrealistically low

yields except at site CS. PC showed an increase in wheat yield ~6.6% in the RCP2.6 and ~15.1% in the RCP6.0 simulations by the end of the 21st century (Fig. 15). Although the different crop models predict different absolute magnitudes of wheat yields, the trend and the interannual variations are similar among all models. Moreover, the magnitude of increase shown by PC is similar to that shown with LPJmL. All models showed increases in wheat yield at individual sites over the 21st century, with the exception of the PEPIC model at the YG site.

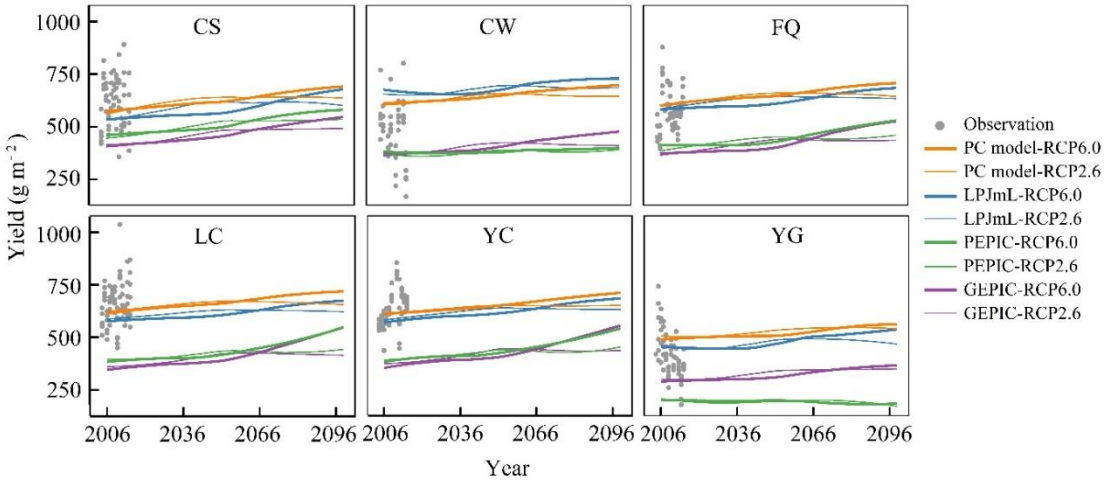


Figure 15: Comparison of different crop models: future scenarios at six sites. The climate data to drive the crop models were derived from the MIROC5 climate model. Lines represent modelled interannual yield trends; Points represent measured yields.

4. Discussion

The P model has been extensively tested against GPP derived from flux measurements in natural vegetation (Stocker et al., 2019; Wang et al., 2017) and has also been shown to perform well in simulating the GPP of croplands (Balzarolo et al., 2018). The present study has confirmed that the P model can predict the GPP of irrigated and fertilized wheat crops in China (Fig. 3); that above-ground crop biomass can be modelled as a fraction of accumulated GPP (Fig. 4); and that yield can be modelled as a saturating function of AB (Fig. 5).

539 We used data from China including several agricultural sites located in major wheat-growing
540 regions to investigate the relationships between AB and yield. These agricultural sites are almost
541 always fertilized and irrigated, creating near-optimal growth conditions. The parameters derived
542 from this study might therefore be appropriate for regions where nutrient and water stresses are
543 absent. Further extensions and tests of the model in a wider range of environmental and economic
544 settings will be needed to allow application to wheat crops that are subject to water and/or nutrient
545 stresses (which are expected to result in different relationships among GPP, AB and yield), or in a
546 wider range of climates.

547 The ratio of above-ground biomass production to GPP is typically 0.41 (natural) to 0.53
548 (managed) for forests, ~0.2 for natural grasslands and 0.6 to 0.7 for managed grasslands (Campioli et
549 al., 2015). Values of this ratio up to ~0.8 have been found for intensively managed crops (Campioli
550 et al., 2015; Chen et al., 2018; Huang et al., 2018). The value of 0.72 estimated in our study is in the
551 upper part of the expected range. This is not unreasonable. As an annual crop, wheat does not require
552 strong roots for support. Modern varieties of wheat are highly efficient in converting GPP to biomass
553 because selective breeding has aimed to increase the GPP allocation to biomass, and ultimately to
554 grain. In addition, the study sites are irrigated (to eliminate water stress) and fertilized (to reduce or
555 eliminate nutrient stress), so the below-ground carbon allocation needed to acquire nutrients and
556 water is minimal.

557 Most crop models assume that grain yield is a fixed proportion of above-ground or total biomass,
558 the so-called harvest index (HI) (Donald, 1962; Hay, 1995). However, the grain yield data analysed
559 here follow a saturation function with AB (Fig. 5), so that HI declines with increasing AB. The level
560 of AB at which saturation occurs is largely determined by the wheat variety (Fig. 6 and Fig. S1). The
561 maximum yield given by ($a \cdot N + b$) in equation (5) is influenced by the wheat variety and the
562 amount of nitrogen added. The actual yield is also determined by the amount of biomass
563 accumulated and, therefore, by the GPP during the growing season – which depends on CO_2, climate,

564 canopy development and incident PPFD. The negative intercept (d) quantifies the requirement for a
565 certain minimum biomass accumulation before any carbon is allocated to seeds. When a linear
566 regression was fitted instead of a saturating function, the estimate of d became positive, which is
567 biologically impossible as it suggests a positive yield when AB is zero (Fig. S2).

568 It follows from this simple empirical representation of the experimental data that improving
569 grain yield is not simply a case of adding more fertilizer (which also comes with significant
570 monetary and environmental costs). Moreover, yields will also not automatically increase in
571 proportion to the effect of CO_2 on photosynthesis, because the saturating nature of this relationship
572 implies a "diminishing return" on increases in AB. The differences among varieties are potentially
573 important here. They suggest that a key target for crop improvement should be the ability of the
574 plants to allocate more carbon to grain as AB increases, and thereby to profit from higher CO_2 levels.

575 Both the current level and the trend in yield over the 21st century simulated by the PC model are
576 similar to that shown by the LPJmL model (Fig. 15). This similarity is probably due to the fact that
577 the LPJ component of LPJmL, is also, ultimately, based on the standard model of photosynthesis and
578 the acclimation of V_{cmax} – the latter process now supported by a wealth of evidence (Smith et al.,
579 2019). However, the PC model is simpler, has fewer parameters and is more transparent, with major
580 advantages both for the credibility of the results and the ease with which uncertainties can be
581 quantified and traced to their source.

582 Quantification of prediction uncertainties in complex models requires extensive computation to
583 estimate the sensitivity to their many parameters. In contrast, the PC model consists of a single
584 central equation (1a), which can readily be differentiated with respect to its (far fewer) parameters.
585 This process allows uncertainties to be attached to predictions without excessive computational
586 demands and allows the major sources of uncertainty to be pinpointed. We have shown that the
587 parameter c^*, related to the metabolic costs of maintaining electron transport capacity, and parameter

588 *b*, related to the potential maximum yield, accounts for a large fraction of the prediction uncertainty
589 (Fig. 10) – indicating the importance of further work to improve these aspects of the model.

590 Studies have suggested that rising temperatures could greatly reduce the grain yield of wheat
591 (Asseng et al., 2014; Asseng et al., 2011; Zhao et al., 2017; Zhao et al., 2016), because of the
592 shortened growing season. However, many studies have neglected the effects of rising CO_2 on C_3
593 photosynthesis, which has the potential to mitigate the impact of rising temperatures on production
594 by improving LUE, particularly as the temperature optimum shifts to higher temperatures with rising
595 CO_2. The effects of elevated CO_2 have been shown by Free Air Carbon dioxide Enrichment (FACE)
596 experiments, with positive impacts on wheat yields and net assimilation rates (Broberg et al., 2019).

597 The primary mechanism by which increasing CO_2 increases GPP is through the improvement of
598 LUE. This mechanism is amplified by the positive feedback by which increased GPP allows greater
599 LAI, which in turn implies greater light absorption and GPP. On the other hand, LMA increases with
600 CO_2, resulting in a diminished response and, according to our model, a peak of the positive response
601 of LAI to CO_2 above ~600 ppm (Fig. 14b and Fig. S4). However, we see no peak in the positive
602 response of yield to CO_2. This appears to be because *f*APAR is comparatively insensitive to changes
603 in LAI at the high end. However, according to our simulations, a positive but saturating response of
604 yield and LUE to CO_2 are found at high CO_2 levels (Fig. 13 and Fig. 14a). This appears to be
605 inconsistent with the yield response shown in the meta-analysis of enhanced CO_2 experiments by
606 Broberg et al. (2019). However, the response shown in that paper is small, solely derived from
607 chamber rather than FACE experiments, and seems to reflect the reduced sensitivity of higher yield
608 wheat varieties to CO_2 changes.

609 The magnitude of yield enhancement simulated by our model is consistent with the findings of
610 Broberg et al. (2019) for the mid-range of wheat yields. Field warming experiments, summarised in
611 Zhao et al. (2016), show negative responses of yield to warming of between 0.5 to 3.0°C at
612 individual sites in northwestern and northern China, with an average response of −4.4% per °C in

613 northwestern China to −2.8% per °C in northern China. Across our study sites (which are in the same
614 region), we predict a net negative response to increased temperature of between −2.3 and −5.7 per °C.
615 This response is caused by the reduction in the length of the growing season and the negative impact
616 of temperature on LUE. The modelled response of yield to combinations of raised CO_2 and
617 temperature, as shown by sensitivity analysis, reflects a combination of net positive effects of CO_2
618 and net negative effects of rising temperature. Scenario runs show that, under the scenarios tested,
619 the positive effects of CO_2 on yield however outweigh the negative effects of temperature, consistent
620 with the findings of other crop models for China (Liu et al., 2019). Warmer regions are more
621 sensitive to warming than cooler regions (also consistent with Liu et al. (2019)), indicating that
622 wheat production in warmer regions of China will be more challenged by climate change.
623
624 **5. Conclusions**
625 The yield of irrigated and fertilized wheat crops at research sites across the wheat-growing
626 region of China was simulated successfully using a parsimonious model based on a combination of
627 first-principles theoretical considerations governing GPP with empirical relationships amongst GPP,
628 AB and yield. The model reproduced the general magnitude and patterns of interannual variability in
629 both AB and yield. When driven by future climate and CO_2 scenarios, it produced results similar to
630 those of the most credible of the more complex crop models.
631 The model also provided insights into how wheat yields may respond to global environmental
632 change. The effect of rising CO_2 on photosynthesis does not imply proportionately increased yield.
633 The model results suggested that the positive response of yield to rising CO_2 may saturate at around
634 600 ppm. The model also predicted a negative effect of warming on wheat yields. Sensitivity
635 analysis showed this negative effect to be stronger in regions with warmer climates today.
636 Nonetheless, in common with other crop models, the simulations indicated an increase of ~6.6% in
637 wheat yields under the RCP2.6 and ~15.1% under the RCP6.0 scenarios of future CO_2 and climate in

China, on the assumption that water and nutrient stresses are eliminated by irrigation and fertilization when required.

Acknowledgements

The authors thank Huimin Lei and Dawen Yang for providing the data at the WeiShan site, and Yiqi Luo, Kun Yang, Chaoqing Yu and Zaichun Zhu for their comments on the analysis. This research was supported by the National Key R&D Program of China (no. 2018YFA0605400). ICP and SPH acknowledge support from the High-End Foreign Expert programme of the China State Administration of Foreign Expert Affairs (no. 161207002). SPH acknowledges funding from the European Research Council (ERC) for "GC2.0: Unlocking the past for a clearer future". This research contributes to the AXA Chair Programme in Biosphere and Climate Impacts and the Imperial College initiative on Grand Challenges in Ecosystems and the Environment (ICP). ICP also acknowledges funding from the ERC, under the European Union's Horizon 2020 research and innovation programme (grant agreement No: 787203 REALM).

References

Ainsworth, E.A. and Rogers, A., 2007. The response of photosynthesis and stomatal conductance to rising CO_2: mechanisms and environmental interactions. Plant Cell Environ, 30(3): 258-270.

Asseng, S. et al., 2014. Rising temperatures reduce global wheat production. Nature Climate Change, 5(2): 143-147.

Asseng, S., Foster, I. and Turner, N.C., 2011. The impact of temperature variability on wheat yields. Global Change Biology, 17(2): 997-1012.

Asseng, S. et al., 2004. Simulated wheat growth affected by rising temperature, increased water deficit and elevated atmospheric CO_2. Field Crops Research, 85(2-3): 85-102.

Balzarolo, M. et al., 2018. TerrA-P: Development and validation of a global GPP/NPP model using MERIS and Sentinel-3 data. Validation report, Available at https://terra-p.vito.be/about/deliverables.

Bernacchi, C.J., Pimentel, C. and Long, S.P., 2003. In vivo temperature response functions of parameters required to model RuBP-limited photosynthesis. Plant Cell Environ, 26(9): 1419-1430.

Bernacchi, C.J., Singsaas, E.L., Pimentel, C., Portis, A.R. and Long, S.P., 2001. Improved temperature response functions for models of Rubisco-limited photosynthesis. Plant Cell Environ, 24(2): 253-259.

Betts, A., Jia, P.W. and Dodson, J., 2014. The origins of wheat in China and potential pathways for its introduction: A review. Quaternary International, 348: 158-168.

Blanc, E., 2017. Statistical emulators of maize, rice, soybean and wheat yields from global gridded crop models. Agricultural and Forest Meteorology, 236: 145-161.

Bondeau, A. et al., 2007. Modelling the role of agriculture for the 20th century global terrestrial carbon balance. Global Change Biology, 13(3): 679-706.

Boylan, M., 2016. What Have We Learned From 15 Years of Supporting the Development of Innovative Teaching Technology? Soc Sci Comput Rev, 22(4): 405-425.

Broberg, M.C., Hogy, P., Feng, Z.Z. and Pleijel, H., 2019. Effects of Elevated CO_2 on Wheat Yield: Non-Linear Response and Relation to Site Productivity. Agronomy-Basel, 9(5): 243.

675 Brooking, I.R., 1996. Temperature response of vernalization in wheat: A developmental analysis. Ann Bot-London, 78(4):
676 507-512.
677 Campbell, G.S. and Norman, J.M., 2012. An introduction to environmental biophysics. Springer Science & Business
678 Media.
679 Campioli, M. et al., 2015. Biomass production efficiency controlled by management in temperate and boreal ecosystems.
680 Nature Geoscience, 8(11): 843-846.
681 Cao, L.J. et al., 2017. Climatic warming in China during 1901-2015 based on an extended dataset of instrumental
682 temperature records. Environmental Research Letters, 12(6).
683 Challinor, A.J. and Wheeler, T.R., 2008. Crop yield reduction in the tropics under climate change: Processes and
684 uncertainties. Agricultural and Forest Meteorology, 148(3): 343-356.
685 Chen, Z., Yu, G.R. and Wang, Q.F., 2018. Ecosystem carbon use efficiency in China: Variation and influence factors.
686 Ecological Indicators, 90: 316-323.
687 Collins, M. et al., 2014. Long-term Climate Change: Projections, Commitments and Irreversibility. Climate Change 2013:
688 The Physical Science Basis: 1029-1136.
689 Donald, C., 1962. In search of yield. J. Aust. Inst. Agric. Sci., 28: 171-178.
690 Elliott, J. et al., 2015. The Global Gridded Crop Model Intercomparison: data and modeling protocols for Phase 1 (v1.0).
691 Geoscientific Model Development, 8(2): 261-277.
692 Fang, H. et al., 2019. Validation of global moderate resolution leaf area index (LAI) products over croplands in
693 northeastern Chinas. Remote Sensing of Environment, 233: 111377.
694 FAOSTAT, 2018. FAOSTAT statistical database, Available at http://www.fao.org/faostat/en/#data/QC.
695 Farquhar, G.D., von Caemmerer, S. and Berry, J.A., 1980. A biochemical model of photosynthetic CO_2 assimilation in
696 leaves of C_3 species. Planta, 149(1): 78-90.
697 Franklin, O. et al., 2017. Using natural selection and optimization for smarter vegetation models - challenges and
698 opportunities, Egu General Assembly Conference.
699 Fu, L., Qu, Y. and Wang, J., 2016. Bias analysis in validation of MODIS LAI product: A case study in cropland of
700 Huailai, northern China, 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS).
701 IEEE, pp. 5921-5924.
702 Gerbaud, A. and Andre, M., 1980. Effect of CO_2, O_2, and Light on Photosynthesis and Photorespiration in Wheat. Plant
703 Physiol, 66(6): 1032-6.
704 Hay, R.K.M., 1995. Harvest Index - a Review of Its Use in Plant-Breeding and Crop Physiology. Annals of Applied
705 Biology, 126(1): 197-216.
706 He, L. et al., 2015. Impacts of recent climate warming, cultivar changes, and crop management on winter wheat
707 phenology across the Loess Plateau of China. Agricultural and Forest Meteorology, 200: 135-143.
708 Huang, J.X. et al., 2016. Assimilating a synthetic Kalman filter leaf area index series into the WOFOST model to
709 improve regional winter wheat yield estimation. Agricultural and Forest Meteorology, 216: 188-202.
710 Huang, K. et al., 2018. Enhanced peak growth of global vegetation and its key mechanisms. Nat Ecol Evol, 2(12): 1897-
711 1905.
712 JCGM, J., 2008. Evaluation of measurement data—Guide to the expression of uncertainty in measurement. Int. Organ.
713 Stand. Geneva ISBN, 50: 134.
714 Kirtman, B. et al., 2014. Near-term Climate Change: Projections and Predictability. Climate Change 2013: The Physical
715 Science Basis: 953-1028.
716 Lawlor, D.W. and Mitchell, R.A.C., 1991. The Effects of Increasing CO_2 on Crop Photosynthesis and Productivity - a
717 Review of Field Studies. Plant Cell Environ, 14(8): 807-818.
718 Lei, H.M. and Yang, D.W., 2010. Seasonal and interannual variations in carbon dioxide exchange over a cropland in the
719 North China Plain. Global Change Biology, 16(11): 2944-2957.
720 Liu, B. et al., 2019. Global wheat production with 1.5 and 2.0 degrees C above pre-industrial warming. Global Change
721 Biology, 25(4): 1428-1444.
722 Liu, J.G., Williams, J.R., Zehnder, A.J.B. and Yang, H., 2007. GEPIC - modelling wheat yield and crop water
723 productivity with high resolution on a global scale. Agr Syst, 94(2): 478-493.
724 Liu, W.F. et al., 2016. Global investigation of impacts of PET methods on simulating crop-water relations for maize.
725 Agricultural and Forest Meteorology, 221: 164-175.
726 Liu, Y.J. et al., 2018. Modelling the impacts of climate change and crop management on phenological trends of spring
727 and winter wheat in China. Agricultural and Forest Meteorology, 248: 518-526.
728 Long, S.P., 1991. Modification of the Response of Photosynthetic Productivity to Rising Temperature by Atmospheric
729 CO_2 Concentrations - Has Its Importance Been Underestimated. Plant Cell Environ, 14(8): 729-739.
730 Maire, V. et al., 2012. The coordination of leaf photosynthesis links C and N fluxes in C_3 plant species. PLoS One, 7(6):
731 e38345.
732 Monsi, M., 1953. Uber den Lichtfaktor in den Pflanzen-gesellschaften und seine Bedeutung fur die Stoffproduktion. Jap.
733 Journ. Bot., 14: 22-52.
734 Muller, C. and Robertson, R.D., 2014. Projecting future crop productivity for global economic modeling. Agricultural

Economics, 45(1): 37-50.
Neitsch, S.L., Arnold, J.G., Kiniry, J.R. and Williams, J.R., 2011. Soil and water assessment tool theoretical documentation version 2009, Texas Water Resources Institute.
Nelson, G.C. et al., 2014. Agriculture and climate change in global scenarios: why don't the models agree. Agricultural Economics, 45(1): 85-101.
Nielsen, D.C. and Halvorson, A.D., 1991. Nitrogen Fertility Influence on Water-Stress and Yield of Winter-Wheat. Agron J, 83(6): 1065-1070.
Ostberg, S., Schewe, J., Childers, K. and Frieler, K., 2018. Changes in crop yields and their variability at different levels of global warming. Earth System Dynamics, 9(2): 479-496.
Palosuo, T. et al., 2011. Simulation of winter wheat yield and its variability in different climates of Europe: A comparison of eight crop growth models. European Journal of Agronomy, 35(3): 103-114.
Pan, H.Z., Chen, Z.X., Ren, J.Q., Li, H. and Wu, S.R., 2019. Modeling winter wheat leaf area index and canopy water content with three different approaches using Sentinel-2 multispectral instrument data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12(2): 482-492.
Piao, S. et al., 2010. The impacts of climate change on water resources and agriculture in China. Nature, 467(7311): 43-51.
Porter, J.R. and Gawith, M., 1999. Temperatures and the growth and development of wheat: a review. European Journal of Agronomy, 10(1): 23-36.
Porter, J.R. et al., 2014. Food Security and Food Production Systems. Climate Change 2014: Impacts, Adaptation, and Vulnerability, Pt A: Global and Sectoral Aspects: 485-533.
Prentice, I.C., Dong, N., Gleason, S.M., Maire, V. and Wright, I.J., 2014. Balancing the costs of carbon gain and water transport: testing a new theoretical framework for plant functional ecology. Ecol Lett, 17(1): 82-91.
Prentice, I.C., Liang, X., Medlyn, B.E. and Wang, Y.P., 2015. Reliable, robust and realistic: the three R's of next-generation land-surface modelling. Atmospheric Chemistry and Physics, 15(10): 5987-6005.
Prentice, I.C. and Thomas, R., 2018. Development and validation of a global GPP/NPP model using MERIS and Sentinel-3 data (TerrA-P) ATBD v2, Available at https://terra-p.vito.be/about/deliverables.
Qin, X.L. et al., 2015. Wheat yield improvements in China: Past trends and future directions. Field Crops Research, 177: 117-124.
Rosenzweig, C. et al., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. Proceedings of the National Academy of Sciences, 111(9): 3268-3273.
Sage, R.F., Sharkey, T.D. and Seemann, J.R., 1989. Acclimation of Photosynthesis to Elevated CO_2 in Five C_3 Species. Plant Physiol, 89(2): 590-6.
Schaphoff, S. et al., 2018. LPJmL4-a dynamic global vegetation model with managed land - Part 1: Model description. Geoscientific Model Development, 11(4): 1343-1375.
Sharp, I., Sanchez-Azofeifa, A. and Musilek, P., 2018. Land Product Validation of MODIS Derived FPAR products over a tropical dry-forest.
Smith, N.G. et al., 2019. Global photosynthetic capacity is optimized to the environment. Ecol Lett, 22(3): 506-517.
Stocker, B.D. et al., 2019. P-model v1. 0: An optimality-based light use efficiency model for simulating ecosystem gross primary production. Geosci. Model Dev. Discuss, 37: 1-59.
Tao, F.L., Zhang, S.A. and Zhang, Z., 2012. Spatiotemporal changes of wheat phenology in China under the effects of temperature, day length and cultivar thermal characteristics. European Journal of Agronomy, 43: 201-212.
Thilakarathne, C.L. et al., 2013. Intraspecific variation in growth and yield response to elevated CO_2 in wheat depends on the differences of leaf mass per unit area. Functional Plant Biology, 40(2): 185-194.
Wang, F.H. et al., 2009. Wheat cropping systems and technologies in China. Field Crops Research, 111(3): 181-188.
Wang, H. et al., 2017. Towards a universal model for carbon dioxide uptake by plants. Nat Plants, 3(9): 734-741.
Wu, J., Gao, X.J., Giorgi, F. and Chen, D.L., 2017. Changes of effective temperature and cold/hot days in late decades over China based on a high resolution gridded observation dataset. International Journal of Climatology, 37: 788-800.
Yang, P. et al., 2007. Evaluation of MODIS land cover and LAI products in cropland of North China Plain using in situ measurements and Landsat TM images. IEEE Transactions on Geoscience and Remote Sensing, 45(10): 3087-3097.
Yu, C. et al., 2019. Managing nitrogen to restore water quality in China. Nature, 567(7749): 516-520.
Zhao, C. et al., 2017. Temperature increase reduces global yields of major crops in four independent estimates. Proc Natl Acad Sci U S A, 114(35): 9326-9331.
Zhao, C. et al., 2016. Field warming experiments shed light on the wheat yield response to temperature in China. Nat Commun, 7: 13530.