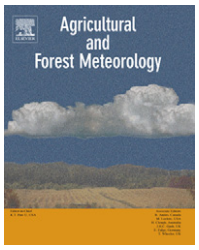
available at www.sciencedirect.com

ScienceDirect

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



Modelling the impacts of weather and climate variability on crop productivity over a large area: A new process-based model development, optimization, and uncertainties analysis

Fulu Tao [a,*], *Masayuki Yokozawa* [b], *Zhao Zhang* [c]

[a] *Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China*
[b] *National Institute for Agro-Environmental Sciences, 3-1-3 Kannondai, Tsukuba, Ibaraki 305-8604, Japan*
[c] *State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China*

ARTICLE INFO

Article history:
Received 28 March 2008
Received in revised form
5 November 2008
Accepted 7 November 2008

Keywords:
Agriculture
Climate change
CO_2 fertilization effects
Transpiration
Water use efficiency
Yield prediction

ABSTRACT

Process-based crop models are increasingly being used to investigate the impacts of weather and climate variability (change) on crop growth and production, especially at a large scale. Crop models that account for the key impact mechanisms of climate variability and are accurate over a large area must be developed. Here, we present a new process-based general Model to capture the Crop–Weather relationship over a Large Area (MCWLA). The MCWLA is optimized and tested for spring maize on the Northeast China Plain and summer maize on the North China Plain, respectively. We apply the Bayesian probability inversion and a Markov chain Monte Carlo (MCMC) technique to the MCWLA to analyze uncertainties in parameter estimation and model prediction and to optimize the model. Ensemble hindcasts (by perturbing model parameters) and deterministic hindcasts (using the optimal parameters set) were carried out and compared with the detrended long-term yields series both at the crop model grid ($0.5^\circ \times 0.5^\circ$) and province scale. Agreement between observed and modelled yield was variable, with correlation coefficients ranging from 0.03 to 0.88 ($p < 0.01$) at the model grid scale and from 0.45 to 0.82 ($p < 0.01$) at the province scale. Ensemble hindcasts captured significantly the interannual variability in crop yield at all the four investigated provinces from 1985 to 2002. MCWLA includes the process-based representation of the coupled CO_2 and H_2O exchanges; its simulations on crop response to elevated CO_2 concentration agree well with the controlled-environment experiments, suggesting its validity also in future climate. We demonstrate that the MCWLA, together with the Bayesian probability inversion and a MCMC technique, is an effective tool to investigate the impacts of climate variability on crop productivity over a large area, as well as the uncertainties.

1. Introduction

In order to establish food security warning systems, predict regional food production in future, and examine the options for adaptations, the impacts of weather and climate variability (change) on crop growth and productivity must be simulated at a large scale. Crop models are increasingly being used on a large spatial scale, often coupled with general circulation

* *Corresponding author.*
E-mail addresses: taofl2002@yahoo.com, taofl@igsnrr.ac.cn (F. Tao).
0168-1923/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.agrformet.2008.11.004

Nomenclature

a	leaf respiration as a fraction of Rubisco capacity
A_{dt}	daytime assimilation rate
A_{gd}	daily gross photosynthesis
A_{nd}	daily leaf net photosynthesis
A_{PAR}	daily integral of absorbed P_{AR}
c_a	ambient mole fraction of CO_2
c_c	carbon content of biomass
c_p	specific heat of moist air
D_{rmax}	crop maximum root depth
E	daily evapotranspiration
E_{demand}	atmospheric demand water
E_{eq}	equilibrium evapotranspiration
E_{supply}	crop- and soil-limited water supply function
E_{vp}	evaporation from the soil evaporation layer
f_{PAR}	fraction of incoming P_{AR} intercepted by green vegetation
f_{temp}	temperature inhibition function limiting photosynthesis at low and high temperatures
$f(z)$	specific root fraction
F	flooding stress factor
F_{cr}	constant to adjust the damage degree of a flooding event
g_c	canopy conductance
g_m	empirical parameter in calculating E_{demand}
g_{min}	minimum canopy conductance
g_{pot}	non-water-stressed potential canopy conductance
h	day length in hours
H_I	harvest index
I_c	precipitation interception storage parameter
I_t	precipitation interception by the leaf canopy
k_b	light extinction coefficient
k_c	kinetic parameter with a Q_{10} dependence on temperature
k_o	kinetic parameter with a Q_{10} dependence on temperature
k_{perc}	soil-texture-dependent percolation rate at field capacity
L_{AI}	leaf area index
L_{AImax}	maximum leaf area index
L_{AIdg}	mean rate of L_{AI} decrease after flowering to maturity
m_c	moisture content of grain
m_r	an empirical parameters in calculating maintenance respiration
M_{lt}	maximum melt rate of the snow pack
p_a	ambient partial pressure of CO_2
p_i	intercellular partial pressure of CO_2 (Pa)
pO_2	ambient partial pressure of O_2 (Pa)
p_{re}	atmospheric pressure
P	daily total precipitation
P_{AR}	photosynthetically active radiation
P_{er}	daily percolation
R_d	daily leaf respiration
R_g	growth respiration
R_m	maintenance respiration
R_{m25}	maintenance respiration at 25 °C
R_n	daily total net radiation flux
$R_{r:l}$	relative growth rate of root depth and leaf area index
s	rate of increase of saturated vapour pressure with temperature
S	soil water stress factor
S_{cr}	critical threshold value of S to affect growth
S_{le}	scaling factor for absorbed P_{AR} at ecosystem versus leaf scale
S_t	precipitation interception storage by leaf canopy
t	time
t_{sec}	number of daylight seconds per day
$\bar{T}$	mean daily temperature
T_b	base temperature
TDD	thermal time
T_{eff}	effective temperature
T_m	maximum temperature
T_o	optimum temperature
T_{snow}	mean daily temperature below which precipitation falls as snow
T_T	transpiration rate
T_{Tmax}	maximum transpiration rate
T_{Tpot}	potential transpiration
V_{EF}	root extraction front velocity
V_m	maximum daily rate of net photosynthesis
V_{PD}	vapour pressure deficit
W	aboveground biomass
W_{ep}	volumetric water content of the evaporation layer, expressed as a fraction of its available water holding capacity
W_o	the initial biomass at emergence
W_s	volumetric water content of the soil layer expressed as a fraction of its available water holding capacity
W_{sow}	threshold fraction of soil water for crop sowing
Y	cumulative fraction of roots between the soil surface and depth z
Yd	crop yield
Y_{gp}	yield gap parameter
z	root depth below soil surface

Greek letters

Γ^*	CO_2 compensation point
α	effective ecosystem-level quantum efficiency
α_g	growth respiration parameter
α_m	empirical parameter in calculating E_{demand}
β	empirical parameter that determines the root distribution with depth
ε	molecular weight ratio of water vapour/dry air
γ	psychrometric constant
λ	latent heat of vaporization of water
λ_i	parameter balancing p_i and p_a
θ	shape parameter that specifies the degree of colimitation by light and Rubisco activity
ρ	density of air
τ	a kinetic parameter with a Q_{10} dependence on temperature
ξ	Priestley–Taylor coefficient

models (e.g., Osborne et al., 2007). However, most dynamic crop models are typically designed to simulate crop growth, yield, and resource utilization at the scale of a homogeneous plot, with relatively high input data requirements. There is a substantial mismatch between spatial and temporal scales of available data and crop simulation models (Hansen and Jones, 2000; Challinor et al., 2004). Different input scales can produce very different simulated yield impacts (Mearns et al., 2001).

To simulate crop growth and productivities over a large area, previous studies adapted process-based models to predict regional yield, such as crop model scaling approaches (Hansen and Jones, 2000) and the yield correction approach (Jagtap and Jones, 2002). Other studies adapted empirical or semi-empirical models with low input data requirements, such as a rice simulation model SIMRIW (Horie et al., 1995), the FAO method (Doorenbos and Kassam, 1979; Martin et al., 2000; Fischer et al., 2002; Tao et al., 2003), and remote-sensing-based production efficiency models (Tao et al., 2005). Challinor et al. (2004) tried to combine the benefits of more empirical modelling methods (low input data requirements, validity over large areas) with the benefits of a process-based approach (the potential to capture variability due to different subseasonal weather patterns and hence increased validity under future climates), resulting in a general large-area model (GLAM) for annual crops. However, like many other crop models, GLAM did not include several key biophysical processes that are important in determining crop response to climate variability, particularly in future climate. For example, there is an need for more process-based modelling of the impact of vapour pressure deficit (V_{PD}), and the combined effects of temperature and elevated CO_2 concentration ([CO_2]) on photosynthesis, transpiration and water use efficiency (Tubiello et al., 2007a).

Extensive controlled-environment experiments such as the Free-Air Concentration Enrichment experiments (e.g., Kimball et al., 1995; Ainsworth et al., 2002; Leakey et al., 2004; Kim et al., 2006, 2007) have showed that increases in both mean and extremes of temperature and elevated [CO_2], under predicted climate change scenarios, can impact the growth and development of crops in several ways. Sustained temperature increases over the season will change the growing period of a crop (e.g., IPCC, 2001), whereas short episodes of high temperature during the critical flowering period of a crop can impact yield independently of any substantial changes in mean temperature (e.g., Matsui and Horie, 1992; Wheeler et al., 2000). Temperature is also a key determinant of evaporative and transpirative demand (e.g., Priestley and Taylor, 1972). Crops sense and respond directly to rising atmospheric CO_2 through increased photosynthesis and reduced stomatal conductance (Jarvis and William, 1998). All other effects of elevated [CO_2] on plants and ecosystems are derived from these two fundamental responses (Long et al., 2004). Rising CO_2 would increase the photosynthesis rate, especially for C_3 crops (Kimball et al., 1995). Although C_4 crops may not show a direct response in photosynthesis activity, an indirect increase in water use efficiency in water-stressed environments via reduction in stomatal conductance may still increase yield (Long et al., 2004). Under elevated CO_2, stomatal conductance in most species will decrease, which may result in less transpiration per unit leaf area (Sionit et al., 1984; Atkinson et al., 1991). Water loss by transpiration is not only affected by the conductivity of the stomata, but also by the driving forces for exchange of the water vapour from the leaf surface to the surrounding atmosphere (i.e., V_{PD}; McNaughton and Jarvis, 1991; Kimball et al., 1995). With all other factors being equal, the existing V_{PD} between stomatal cavity and surrounding air – the boundary layer – will increase at a reduced transpiration rate and feedback to stimulate transpiration.

Although most dynamic global vegetation models have accounted for such key response mechanisms by coupling photosynthesis and stomatal conductance (Cramer et al., 2001), many crop models often simulate the key responses of crop to climate change (such as CO_2 fertilization effects and change in transpiration) using a proportionality factor (Long et al., 2006; Tubiello et al., 2007b). The important consideration is that experimentally observed crop physiological responses to climate change variables at plot and field levels (e.g. Kimball et al., 1995; Ainsworth et al., 2002; Leakey et al., 2004; Kim et al., 2006, 2007) are too simplified in current crop models (Tubiello et al., 2007a). As a consequence, the potential for negative surprises is not fully explored, thus reducing the level of confidence in regional and global projections (Tubiello et al., 2007a). It is thus imperative to continue to advance the fundamental knowledge of crop species responses to climate change, reduce uncertainties in impact projections, and assess future risks (Tubiello et al., 2007a).

Here, we develop a new process-based Model to capture the Crop–Weather relationship over a Large Area (MCWLA). The MCWLA is designed to investigate the impacts of weather and climate variability on crop growth and productivity at a large scale. Toward this aim, we tried to capture the interannual variability in observed crop yield and water use by accounting for subseasonal variability in weather and crop responses. Most importantly, the MCWLA also simulates crop response to elevated [CO_2] and high temperature by adopting photosynthesis–stomatal conductance coupling. In the meantime, like GLAM (Challinor et al., 2004), the impacts on yield due to factors other than weather (e.g., pests, disease, management factors) are modelled in a simplified way.

We apply the Bayesian probability inversion and a Markov chain Monte Carlo (MCMC) technique to the MCWLA to analyze uncertainties in parameter estimation and model prediction and to optimize the model. Ensemble hindcasts (by perturbing model parameters) and deterministic hindcasts (using the optimal parameters set) were carried out and compared with the detrended long-term yields series both at the crop model grid and the province scale.

The MCWLA is a general crop model. In this study, the model is optimized and tested for spring maize in the Heilongjiang and Jilin provinces on the Northeast China Plain and summer maize in the Henan and Shandong provinces on the North China Plain, respectively (Fig. 1). Maize (*Zea mays*) is the most widely cultivated C_4 crop ranking as the third most produced food crop in China and the world. Any effects of increasing temperature and elevated [CO_2] on maize are likely to have significant consequences in terms of global food production (Leakey et al., 2004; Tao et al., 2008a). Extension to

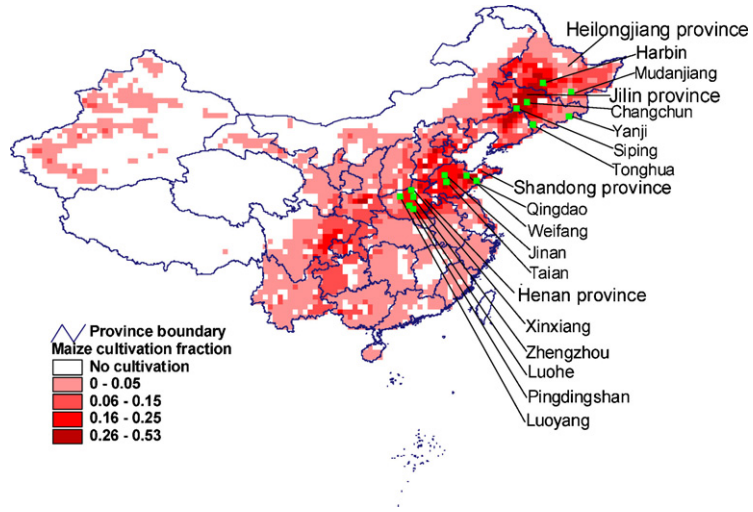


Fig. 1 – Maize cultivation fraction in China at 0.5° × 0.5° grid resolution and the provinces and grids analyzed in this study.

other crop and/or regions can proceed along similar lines to the calibration described in Section 3.

2. Model description

2.1. Growth and development

MCWLA simulates crop growth and development in a daily time-step. As in most crop models, growing degree-days provide the driving force for the processes of canopy development, flowering, and maturity. As Challinor et al. (2004), the crop is planted either on a specified date or on the first day that the soil moisture exceeds a given fraction of the maximum available soil water (W_{sow}) within a sowing window. If the threshold is not reached within the sowing window (20 days in this study) then the crop is sowed regardless. From the planting date (pd), the thermal time (TDD_i) elapsed after a given development stage i is given by

$$TDD_i = \sum_{t=pd}^{t_i} (T_{eff} - T_b) \quad (1)$$

where t is the time, T_{eff} is the effective temperature, T_b is the base temperature below which development ceases, and i is the development stage number (equal to 0 from sowing to emergence, 1 from emergence to the beginning floral initiation, 2 from the beginning floral initiation to the end floral initiation, 3 from the end floral initiation to flowering, and 4 from flowering to maturity). Development stage i completes after a specified duration TDD_i has elapsed; and harvest occurs at maturity.

As in GLAM (Challinor et al., 2004), the effective temperature, T_{eff}, is defined as follows using cardinal temperatures T_b, T_o, and T_m, where the subscripts denote base, optimum, and maximum temperature, respectively:

$$T_{eff} = \begin{cases} \bar{T} & T_b \le \bar{T} \le T_o \\ T_o - (T_o - T_b)\left(\dfrac{\bar{T} - T_o}{T_m - T_o}\right) & T_o < \bar{T} < T_m \\ T_b & T \ge T_m,\ \ \bar{T} < T_b \end{cases} \quad (2)$$

where $\bar{T}$ denotes mean daily temperature.

Previous crop modelling studies have suggested the expansion of leaf area be modelled independently of leaf biomass (Horie et al., 1995; Jamieson and Semenov, 2000; Bannayan et al., 2005). In MCWLA, the growth of the crop leaf area is determined as follows, which is improved from the GLAM (Challinor et al., 2004):

$$\frac{\partial L_{AI}}{\partial t} = \begin{cases} \left(\dfrac{T_{eff} - T_b}{TDD_3 - TDD_0}\right) L_{AImax} Y_{gp} \min\left(\dfrac{S}{S_{cr}}, 1\right) \min\left(\dfrac{F}{F_{cr}}, 1\right) & i \le 3 \\ \left(\dfrac{\partial L_{AIdg}}{\partial t}\right) \max\left(\left(1 + \left(1 - \dfrac{S}{S_{cr}}\right)\right), 1\right) \max\left(\left(1 + \dfrac{F}{F_{cr}}\right), 1\right) & i > 3 \end{cases} \quad (3)$$

where L_{AImax} is the maximum leaf area index (L_{AI}) of the crop. The soil water stress factor, S, is

$$S = \frac{T_T}{T_{Tpot}} \quad (4)$$

which begins to affect growth at values less than the critical threshold value S_{cr}, T_T and T_{Tpot} are the rates of transpiration and potential transpiration, respectively. F is the flooding stress factor; its value increases by 1.0 when one flooding event occurs (defined as soil water being above soil water capacity for 3 continuous days) from sowing to maturity. F_{cr} is a parameter to adjust the damage extent of one flooding event. L_{AIdg} is the mean rate of L_{AI} decrease after flowering to maturity. Y_{gp} is the yield gap parameter, used to

reduce L_{AI} from the physical value to an effective value, which accounts for the mean effects of pests, diseases, and non-optimal management, as in the GLAM (Challinor et al., 2004).

The roots grow according to the following equations:

$$\frac{\partial V_{EF}}{\partial L_{AI}} = R_{r:l}\frac{D_{rmax}}{L_{AImax}} \quad (5)$$

where V_{EF} is the extraction front velocity, D_{rmax} is the crop-specific maximum root depth, and $R_{r:l}$ is a parameter to describe the relative growth rate of root depth and L_{AI}. The specific root fraction $f(z)$ is derived from an asymptotic root distribution proposed by Gale and Grigal (1987):

$$Y = 1 - \beta^Z \quad (6)$$

where Y is the cumulative fraction of roots between the soil surface and depth z (cm). β is an empirical fitting parameter that determines the root distribution with depth. A higher β value gives rise to a larger proportion of roots at deeper depths relative to low β values. The specific root fraction function $f(z)$ is the derivative of Eq. (6) with respect to soil depth z and is expressed as

$$f(z) = \frac{\partial Y}{\partial z} = \frac{\partial}{\partial z}(1 - \beta^z) = -\beta^2 In\beta \quad (7)$$

In practice, the β value is estimated from the rooting depth z (cm) after Li et al. (2006):

$$\beta = 0.01\frac{1}{z} \quad (8)$$

Eq. (8) is derived based on the assumption that the total root fraction from the soil surface to the rooting depth z is 0.99 [because Eq. (6) is asymptotic, the β value cannot be derived if the total root fraction is exactly 1.0].

2.2. Soil water balance

Generally, soil hydrology is modelled following the semi-empirical approach of Haxeltine and Prentice (1996a), which was simplified from the model developed by Neilson (1995). In the MCWLA, the soil profile is split into 12 soil layers with a fixed thickness of 15 cm. The water content of each layer is updated daily taking into account snowmelt, percolation, rainfall, evapotranspiration, and runoff. Precipitation falls as rain or snow depending on whether the daily air temperature is above or below T_{snow} (−2 °C). Above this threshold the snow pack begins to melt at a maximum rate of

$$M_{lt} = (1.5 + 0.007P)(\bar{T} - T_{snow}) \quad (9)$$

where P is daily total precipitation.

Precipitation interception (I_t) by the leaf canopy is estimated as Kergoat (1998):

$$I_t = E_{eq}\xi \min\left[\left(\frac{S_t}{E_{eq}\xi}\right), 0.99\right] \quad (10)$$

where E_{eq} is the equilibrium evapotranspiration, ξ is Priestley–Taylor coefficient, and S_t is the interception storage by the leaf canopy, estimated by

$$S_t = \min\lfloor P, (I_c L_{AI} P)\rfloor \quad (11)$$

where I_c is the interception storage parameter. Experimental results from several sites around the world, including vegetated surfaces and large water bodies (lake and oceans), gave ξ values in the range of 1.08 ± 0.01 to 1.34 ± 0.05, with an average of 1.26 (Priestley and Taylor, 1972).

Evaporation from the soil evaporation layer (defined as the upper 20 cm of soil profile), E_{vp}, is estimated as in the CERES models (Ritchie et al., 1988):

$$E_{vp} = \begin{cases} E_{eq}\xi W_{ep}(1 - 0.43L_{AI}) & L_{AI} < 1 \\ E_{eq} W_{ep}\exp(-0.4L_{AI}) & L_{AI} \geq 1 \end{cases} \quad (12)$$

where W_{ep} is the volumetric water content of the evaporation layer, expressed as a fraction of its available water holding capacity.

Daily percolation (P_{er}) from one soil layer to the next is calculated using the empirical relationship of Neilson (1995):

$$P_{er} = k_{perc} W_s^2 \quad (13)$$

where k_{perc} represents the soil texture dependent percolation rate (mm d^{-1}) at field capacity and W_s is the volumetric water content of the soil layer expressed as a fraction of its available water holding capacity. Surface runoff and drainage are calculated as the excess water above field capacity in the first layer and all other layers, respectively.

Daily evapotranspiration (E) is calculated as the minimum of a crop- and soil-limited supply function (E_{supply}) and the atmospheric demand (E_{demand}):

$$E = \min(E_{supply}, E_{demand}) \quad (14)$$

where E_{supply} is the product of crop-root-weighted soil moisture availability and a maximum transpiration rate, T_{Tmax}. The percentage of water extracted by crop roots at the upper, second, third, and bottom quarter of the root zone follows a 40/30/20/10 per cent water extraction pattern (SCS, 1991). E_{demand} is calculated following Monteith's empirical relation between evaporation efficiency and surface conductance (Monteith, 1995; Haxeltine and Prentice, 1996a):

$$E_{demand} = E_{eq}\alpha_m\left[1 - \exp\left(\frac{-g_{pot}}{g_m}\right)\right] \quad (15)$$

where g_{pot} is the non-water-stressed potential canopy conductance calculated by the photosynthesis routine, and g_m and α_m are empirical parameters (Monteith, 1995). E_{eq} is calculated from latitude, temperature, and fractional sunshine hours, using a standard method based on the Prescott equation (Jarvis and MacNaughton, 1986; Prentice et al., 1993):

$$E_{eq} = \frac{[s/(s+\gamma)]R_n}{\lambda} \quad (16)$$

where R_n is the daily total net radiation flux (MJ m^{-2} d^{-1}), γ is a psychrometric constant (kPa °C^{-1}), λ is the latent heat of vaporization of water (MJ kg^{-1}), and s is the rate of increase of saturated vapour pressure with temperature (kPa °C^{-1}):

$$\gamma = \frac{c_p p_{re}}{\varepsilon\lambda} \times 10^{-3} \quad (17)$$

$$\lambda = 2.501 - (2.361 \times 10^{-3})\bar{T} \quad (18)$$

$$s = 2.5 \times 10^3 \frac{\exp[17.27\bar{T}/(237.3+\bar{T})]}{(237.3+\bar{T})^2} \quad (19)$$

where c_p is the specific heat moist air at constant pressure (kJ kg^{-1} °C^{-1}), p_{re} is atmospheric pressure (kPa), and ε is the molecular weight ratio of water vapour/dry air.

2.3. *Photosynthesis–stomatal conductance coupling and transpiration*

In the MCWLA, we use the robust, process-based representation of the coupled CO_2 and H_2O exchanges in the Lund–Postdam–Jena (LPJ) dynamic global vegetation models (Haxeltine and Prentice, 1996a,b; Sitch et al., 2003), which was later used for agriculture (Bondeau et al., 2007). The Farquhar photosynthesis model (Farquhar et al., 1980; Farquhar and von Caemmerer, 1982), as generalized for global modelling purposes by Collatz et al. (1991, 1992), underlies the model. The strong optimality hypothesis (Dewar, 1996; Haxeltine and Prentice, 1996b; Prentice et al., 2000) is assumed to apply; the nitrogen content and Rubisco activity of leaves are assumed to vary both seasonally and with canopy position so as to maximize net assimilation at the leaf level.

The daily integral of absorbed photosynthetically active radiation (P_{AR}), A_{PAR}, is calculated following Haxeltine and Prentice (1996a):

$$A_{PAR} = P_{AR} f_{PAR} S_{le} \quad (20)$$

where S_{le} is a scaling factor for absorbed P_{AR} at the ecosystem versus leaf scale; f_{PAR} is the fraction of incoming P_{AR} intercepted by green vegetation and is estimated by

$$f_{PAR} = 1 - \exp(-k_b L_{AI}) \quad (21)$$

where k_b is a light extinction coefficient.

For C_3 plant assimilation, daily gross photosynthesis, A_{gd} (g C m^{-2} d^{-1}), is given by

$$A_{gd} = A_{PAR} c_1 [1 - \sigma_c] \quad (22)$$

Daily leaf net photosynthesis, A_{nd} (g C m^{-2} d^{-1}), is given by

$$A_{nd} = A_{PAR}\left(\frac{c_1}{c_2}\right)[c_2 - (2\theta - 1)s - 2(c_2 - \theta s)\sigma_c] \quad (23)$$

where θ is a shape parameter that specifies the degree of colimitation by light and Rubisco activity (Haxeltine and Prentice, 1996a,b). The terms σ_c, s, c_1, and c_2 are given by

$$\sigma_c = \left[1 - \frac{c_2 - s}{c_2 - \theta s}\right]^{0.5} \quad (24)$$

$$s = \left(\frac{24}{h}\right) a \quad (25)$$

$$c_1 = \alpha f_{temp} \frac{p_i - \Gamma_*}{p_i + 2\Gamma_*} \quad (26)$$

$$c_2 = \frac{p_i - \Gamma_*}{p_i + k_c(1 + pO_2/k_o)} \quad (27)$$

where h is the day length in hours, a is a constant (leaf respiration as a fraction of Rubisco capacity), α is the effective ecosystem-level quantum efficiency, and f_{temp} is a temperature inhibition function limiting photosynthesis at low and high temperatures (Larcher, 1983). Γ_* is the CO_2 compensation point given by

$$\Gamma_* = \frac{pO_2}{2\tau} \quad (28)$$

where pO_2 is the ambient partial pressure of O_2 (Pa). p_i is the intercellular partial pressure of CO_2 (Pa), given by

$$p_i = \lambda_i p_a \quad (29)$$

where p_a is the ambient partial pressure of CO_2 and λ_i is a parameter. Parameters τ, k_o, and k_c are kinetic parameters with a Q_{10} dependence on temperature (Brooks and Farquhar, 1985; Collatz et al., 1991).

An appropriate simplification of the model (with different values for a and α and saturating p_i) is applied for plants with C_4 physiology (Haxeltine and Prentice, 1996a). Eqs. (23)–(39) describe the biochemical dependence of total daily net assimilation on p_i and environmental variables.

The daytime assimilation rate A_{dt} is also related to p_i through the CO_2 diffusion gradient between the atmosphere and intercellular air spaces:

$$g_c = g_{min} + \frac{1.6A_{dt}}{c_a(1 - \lambda_i)} \quad (30)$$

where g_{min} is the minimum canopy conductance and c_a is the ambient mole fraction of CO_2 ($p_a = p_{re} \cdot c_a$). A_{dt} is obtained from A_{nd} by addition of nighttime respiration:

$$A_{dt} = A_{nd} + \left(\frac{1-h}{24}\right) R_d \quad (31)$$

where R_d is daily leaf respiration in g C m^{-2} d^{-1}, and scaled to V_m, the maximum daily rate of net photosynthesis, by

$$R_d = aV_m \quad (32)$$

The optimal value for V_m is calculated by optimizing Eq. (23) using the constraint $\partial A_{nd}/\partial V_m = 0$ resulting in the following equation for V_m (g C m^{-2} d^{-1}):

$$V_m = \left(\frac{1}{a}\right)\left(\frac{c_1}{c_2}\right)[(2\theta - 1)s - (2\theta s - c_2)\sigma_c]A_{PAR} \quad (33)$$

Under non-water-stressed conditions, maximum values of λ_i are assumed; A_{nd} is calculated from Eq. (23) and g_c is derived from Eq. (30). The value for canopy conductance thus obtained is the potential canopy conductance, g_{pot}, required to derive

demand-limited E in Eq. (15). If water supply limits transpiration, Eqs. (15), (23) and (30) are solved simultaneously to yield values of λ_i and g_c consistent with the transpiration rate.

By assuming the leaf surface temperature is equal to surface atmospheric temperature, as Sellers et al. (1996), the photosynthesis-related T_{Tpot} and T_T are calculated as

$$T_{Tpot} = \frac{t_{sec}(g_{pot} - g_{min})V_{PD}\rho c_p}{\gamma} \quad (34)$$

$$T_T = \frac{t_{sec}(g_c - g_{min})V_{PD}\rho c_p}{\gamma} \quad (35)$$

where t_{sec} is the number of daylight seconds per day and ρ is the density (kg m^{-3}) of air.

2.4. Biomass accumulation and yield formation

Biomass (W in g C m^{-2}) increases from the initial biomass at emergence (W_o) is determined by

$$\frac{\partial W}{\partial t} = A_{gd} - R_m - R_g \quad (36)$$

where R_m is maintenance respiration and R_g is growth respiration. Following (Hunt, 1994; Tao et al., 2005), R_m are given by

$$R_m = R_{m25}\left(\frac{W}{W + m_r}\right)Q_{10}^{(\bar{T}-25)/10} \quad (37)$$

where R_{m25} is the maintenance respiration at 25 °C; m_r is an empirical parameters; temperature dependent Q_{10} for maintenance respiration is modelled as a function of temperature following Tjoelker et al. (2001) as

$$Q_{10} = 3.22 - 0.046\bar{T} \quad (38)$$

R_g is given by

$$R_g = \max(\alpha_g(A_{gd} - R_m), 0) \quad (39)$$

where α_g is growth respiration parameter.

Biomass is transferred into yield, Yd (g m^{-2}), as Lobell et al. (2002), using:

$$Yd = \frac{W}{1 - m_c}c_c H_I \quad (40)$$

where m_c is the moisture content of grain, c_c is carbon content of biomass, and H_I is the harvest index. As in the GLAM (Challinor et al., 2004), for $i \leq 3$ $H_I = 0$, and for $i > 3$

$$\frac{\partial H_I}{\partial t} = \text{constant} \quad (41)$$

3. Parameter calibration, uncertainties, and optimization

3.1. Method

The Bayesian probability inversion and an MCMC technique have been demonstrated as an effective method to synthesize information from various sources for analyzing model uncertainties and optimizing model parameters (Knorr and Kattge, 2005; Xu et al., 2006; Iizumi et al., in press). Here the technique was applied to the MCWLA to analyze uncertainties of parameters and simulated crop yields. We calibrated the MCWLA for spring maize at the 0.5° × 0.5° grid of Harbin (Fig. 1) using the statistical datasets of phenology (planting date, flowering date and maturity date) and yields from 1985 to 1996. Likewise, we calibrated the MCWLA for summer maize at the 0.5° × 0.5° grid of Zhengzhou (Fig. 1) using the observed datasets of phenology and yields from 1995 to 2002.

3.2. Datasets

The MCWLA requires daily weather inputs for mean temperature, precipitation, vapour pressure, and fractional sunshine hours. In this study, the MCWLA was run at each 0.5° × 0.5° grid with maize cultivation fraction ≥0.05 across four major production provinces: Heilongjiang, Jilin, Henan, and Shandong (Fig. 1). Monthly data on mean temperature, vapour pressure, and sunshine hours for the 0.5° × 0.5° resolution grids were obtained from the climate research unit in University of East Anglia, U.K. (Mitchell and Jones, 2005). The monthly means were interpolated to daily values using spline interpolation (Press et al., 1992). Daily precipitation at 0.5° × 0.5° resolution grids was obtained from the APHRODITE project (Asian Precipitation–Highly-Resolved Observational Data Integration Towards Evaluation of the Water Resources), which develops state-of-the-art daily precipitation datasets with high-resolution grids (0.25° and 0.5°) for Asia. The datasets were developed primarily with data obtained from a rain-gauge observation network (Xie et al., 2007).

Soil texture and hydrological properties data were based on the FAO soil dataset (Zobler, 1986; FAO, 1991), as in LPJ dynamic global vegetation models (Sitch et al., 2003). Soil parameters include the soil-texture-dependent percolation rate (mm d^{-1}) at field capacity (k_{perc}) and available volumetric water holding capacity (i.e., the water holding capacity at field capacity minus water holding capacity at the wilting point, expressed as a fraction of soil layer depth).

Yearly district-, county-, or subprovince-level (usually including five to eight counties) data on maize yield and growing area were obtained from the statistical yearbook of each county or province. Yearly maize phenology at the grids of Harbin and Zhengzhou, including planting, flowering, and harvest dates, were obtained from the agricultural meteorological stations in Harbin and Zhengzhou (Tao et al., 2006). Yearly growing-area-weighted yields at some 0.5° × 0.5° grids (Fig. 1) were calculated from their district-level data on growing area and yield. Yearly growing-area-weighted yields for the Heilongjiang, Jilin, Henan, and Shandong provinces (Fig. 1) were calculated from the county- or subprovince-level data on growing area and yield. The growing-area-weighted yields at the 0.5° × 0.5° grids and provinces were detrended to produce yield data at the production technology of the base year, and these data (referred to as 'observed yields') were used in the model calibration and evaluation procedure.

Table 1 – Selected model parameters prior intervals, 97.5% high-probability intervals (lower limit, upper limit), mean estimates, standard deviation, and the optimal parameter set at the grid Harbin for spring maize [Zhengzhou for summer maize].

Parameters	Prior interval	97.5% high-probability interval	Mean	Standard deviations	Parameter value in the optimal set
Phenological parameters					
T_b (°C)	5–15[5–15]	5.9–9.5[7.9–10.0]	7.7[8.9]	1.2[0.6]	9.6[8.9]
T_o (°C)	20–31[20–31]	23.0–30.8[27.7–30.9]	26.9[29.5]	2.5[0.9]	23.3[30.2]
T_m (°C)	31–36[31–36]	31.1–35.9[31.1–35.9]	33.6[33.6]	1.5[1.4]	35.0[35.3]
TDD_0 (degree-days)	50–200[80–200]	57.8–197.9[83.3–197.6]	131.7[147.6]	42.3[33.8]	151.9[172.0]
TDD_1 (degree-days)	200–800[200–600]	357.3–786.9[273.6–581.5]	618.2[455.0]	124.6[86.4]	669.1[587.5]
TDD_2 (degree-days)	500–1000[600–900]	504.7–786.9[603.1–792.9]	627.6[676.7]	81.8[56.8]	631.3[644.3]
TDD_3 (degree-days)	700–1200[800–1200]	803.7–1113.7[928.4–1059.5]	948.8[994.1]	100.3[34.6]	796.7[973.0]
TDD_4 (degree-days)	1200–1800[1400–1800]	1330.0–1791.3[1590.4–1788.7]	1560.1[1695.0]	153.0[55.3]	1304.5[1710.9]
Light, water use, and yield formation parameters					
Y_{gp}	0.2–1.0[0.2–1.0]	0.26–0.99[0.30–0.98]	0.66[0.71]	0.21[0.19]	0.32[0.45]
$\frac{\partial H_I}{\partial t}$	0.002–0.02[0.002–0.02]	0.004–0.009[0.009–0.019]	0.007[0.015]	0.002[0.003]	0.008[0.018]
$R_{r:l}$	0.2–5.0[0.2–5.0]	0.43–4.93[0.86–4.92]	2.97[3.11]	1.36[1.16]	0.32[3.99]
S_{cr}	0.2–0.8[0.2–0.8]	0.22–0.78[0.21–0.79]	0.47[0.49]	0.17[0.17]	0.26[0.54]
S_{le}	0.2–0.8[0.2–0.8]	0.25–0.79[0.28–0.73]	0.52[0.44]	0.15[0.12]	0.46[0.47]
α	0.033–0.073[0.033–0.073]	0.03–0.07[0.03–0.07]	0.05[0.05]	0.01[0.01]	0.058[0.04]
T_{Tmax} (mm m^{-2} d^{-1})	3.0–15.0[3.0–15.0]	4.37–14.80[5.82–14.88]	10.31[11.55]	3.03[2.52]	6.42[14.39]
g_m	2.0–10.0[2.0–10.0]	5.15–9.93[6.73–9.99]	8.07[9.18]	1.33[0.86]	9.64[9.24]
λ_i	0.2–0.6[0.2–0.6]	0.21–0.58[0.21–0.54]	0.39[0.31]	0.12[0.09]	0.21[0.30]
R_{m25} (g C m^{-2} $d^{-1)}$	0.2–0.9[0.2–0.9]	0.21–0.89[0.21–0.85]	0.55[0.48]	0.20[0.19]	0.38[0.22]
m_r (g C m^{-2})	10.0–100.0[10.0–100.0]	12.46–98.38[13.56–98.24]	57.56[61.91]	25.88[25.28]	17.10[69.30]
α_g	0.1–0.5[0.1–0.5]	0.11–0.49[0.10–0.45]	0.31[0.26]	0.11[0.10]	0.15[0.34]
W_o (g C m^{-2})	0.01–0.2[0.01–0.2]	0.01–0.19[0.01–0.19]	0.11[0.10]	0.05[0.06]	0.19[0.13]

3.3. *Application of Bayes' theorem*

A general description of the Bayesian probabilistic inversion is given by Bayes' theorem (e.g., Tarantola, 1987; Leonard and Hsu, 1999; Gill, 2002) in the form of

$$p(c/Z) = \frac{p(Z/c)\,p(c)}{p(Z)} \tag{42}$$

where $p(c)$ is the prior probability density function (PDF) representing prior knowledge about parameter c, $p(Z/c)$ is the conditional probability density of observations Z on c (also called the likelihood function of parameter c), $p(Z)$ is the probability of observations Z, and $p(c/Z)$ is the posterior probability density function (PPDF) of parameter c. The theorem states that the posterior information of model parameter c represented by $p(c/Z)$ can be obtained from the prior information represented by $p(c)$ and the observed information given by $p(Z/c)$. $p(c/Z)$ is often written in the following form:

$$p(c/Z) \propto p(Z/c)\,p(c) \tag{43}$$

that is, $p(c/Z)$ is proportional to $p(Z/c)p(c)$.

From the Bayesian viewpoint, $p(c/Z)$ represents the solution to an inverse problem because it gives a probabilistic description of parameter c over the parameter space. In the context of this study, the PPDF $p(c/Z)$ of model parameter c can be obtained from prior knowledge of parameter c represented by a prior PDF $p(c)$ and information contained in the datasets of historical crop phenology and yields series represented by a likelihood function $p(Z/c)$. To apply Bayes' theorem, after Xu et al. (2006), we first specified the prior PDF $p(c)$ by giving a set of limiting intervals for parameter c, then constructed the likelihood function $p(Z/c)$based on the assumption that errors in the observed data followed Gaussian distributions.

We select the model parameters important for crop phenology, water use, and yield formation (Table 1). The prior PDF $p(c)$ of parameters was specified as a uniform distribution over the intervals as shown in Table 1. These limits are our prior knowledge about the approximate ranges of the parameters. Better prior knowledge on the parameters should result in more accurate estimates; otherwise we would rather use the weak limits to be more objective and general. We assume a uniform distribution $p(c)$ for parameter c with an emphasis on the equal probability of all parameter values occurring within the limits. This may be the best prior to choose in the absence of any other knowledge regarding parameter distributions.

The likelihood function was specified according to distributions of observation errors. Error $e(t)$ in each observation $Z(t)$ at time t is expressed by

$$e(t) = Z(t) - X(t) \tag{44}$$

where $X(t)$ is the modelled value. For the three datasets used in the study (i.e., yearly observations of flowering date, maturity date, and yield), $e(t)$ is expanded as:

$$e(t) = [e_1(t), e_2(t), e_3(t)]^T \tag{45}$$

Corresponding to each modelled variable, there is one random error component $e_i(t) = Z_i(t) - X_i(t)$. We assumed that $e(t)$ followed a multivariate Gaussian distribution with a zero mean. This assumption is commonly made in many studies (Braswell et al., 2005; Raupach et al., 2005), mostly because a

Gaussian distribution in general can approximate errors of various sources well due to the central limit theorem (Von Mises, 1964). With the Gaussian distribution, the PDF of $e(t)$ at time t is given by

$$P(e(t)) \propto \exp\left\{-\frac{1}{2}[Z(t)-X(t)]^T \mathrm{cov}(e_t)^{-1}[Z(t)-X(t)]\right\} \quad (46)$$

where $\mathrm{cov}(e_t)$ is a covariance matrix of vector $e(t)$. With the assumption that each component $e(t)$ is independently and identically distributed over the observation times, the likelihood function $p(Z/c)$ is then the product of the distributions of $e_i(t)$, $i = 1, 3$ (Eq. (46)) at all observation times:

$$P(Z/c) \propto \exp\left\{-\sum_{i=1}^{3}\frac{1}{2\sigma_i^2}\sum_{t\in \mathrm{obs}(Z_i)}[Z_i(t)-X(t)]^2\right\} \quad (47)$$

where constants σ_1^2, σ_2^2, and σ_3^2 are the error variances of flowering date, maturity date, and yield, respectively. Then, with Bayes' theorem, the PPDF of parameter c is given by Eq. (43).

3.4. *Sampling with the Metropolis–Hastings algorithm*

The Metropolis–Hastings (M–H) algorithm is an MCMC technique revealing high-dimensional PDFs of random variables via a sampling procedure (Metropolis et al., 1953; Hastings, 1970; Geman and Geman, 1984; Gelfand and Smith, 1990). To generate a Markov chain in the parameter space, we ran the M–H algorithm by repeating two steps: a proposing step and a moving step, after Xu et al. (2006). In each proposing step, the algorithm generates a new point c^{new} on the basis of the previously accepted point $c^{(k-1)}$ with a proposal distribution $q(c^{new}/c^{(k-1)})$. In each moving step, point c^{new} is tested against the Metropolis criterion to examine if it should be accepted or rejected. For simplicity of notation, we denote $L(c)$ as the targeted stationary distribution $p(c/Z)$. A computer implementation of the M–H algorithm consists the following steps (Spall, 2003):

Step 1: Choose an arbitrary initial point $c^{(0)}$ in the parameter space.
Step 2: (Proposing step). Propose a candidate point c^{new} according to a proposal distribution $q(c^{new}/c^{(k-1)})$.
Step 3: (Moving step). Calculate $P(c^{(k-1)}, c^{new}) = \min\{1, (L(c^{new})q(c^{(k-1)}/c^{new}))/(L(c^{(k-1)})q(c^{new}/c^{(k-1)}))\}$, and compare the value with a random number U from the uniform distribution U [0,1] that is defined on interval [0,1]. Set $c^{(k)} = c^{new}$ if $U \le P(c^{(k-1)}, c^{new})$; otherwise set $c^{(k)} = c^{(k-1)}$. This test criterion is also called the Metropolis criterion.
Step 4: Repeat steps 2 and 3 until enough samples are obtained.

The proposal distribution $q(c^{new}/c^{(k-1)})$ can strongly affect the efficiency of the M–H algorithm. To find an effective proposal distribution, we first made a test run of the algorithm with 60,000 simulations, using a uniform proposal distribution centred at the currently accepted point:

$$C^{new} = C^{(k-1)} + \lfloor r_d(L_m^u - L_m^l) + L_m^l \rfloor \quad (48)$$

where r_d is a random number uniformly distributed between 0 and 1 and L_m^l and L_m^u are the upper and lower values controlling the proposing step size. Based on the test run, we constructed a Gaussian distribution $N(0, \mathrm{cov}^0(c))$, where $\mathrm{cov}^0(c)$ is a diagonal matrix with its diagonal being set to the estimated variances of the parameter c from the initial test run and zero elsewhere. Next we adopted the following proposal distribution to formally execute the consecutive MCMC simulations:

$$c^{new} = c^{k-1} + N[0, \mathrm{cov}^0(c)] \quad (49)$$

In each proposing step of the M–H algorithm a new point c^{new} is generated from its predecessor $c^{(k-1)}$ from a Gaussian distribution with mean $c^{(k-1)}$, constant variances estimated from the previous run, and zero parameter covariance.

We formally made three parallel runs of the M–H algorithm with the proposal distribution in Eq. (49). Each run simulated 60,000 times. The initial number of samples in the burn-in period (5000 samples) was discarded after the running mean and standard deviations were stabilized. The acceptance rates for the newly generated samples were about 30–40% for the three runs. For statistical analysis of the parameters, we used the samples of the final run (55,000 samples in total) after their burn-in period.

3.5. *Parameter estimation*

We estimated parameter statistics based on the 55,000 samples of the final run. Uncertainties of the parameters were quantified with a 97.5% highest probability density interval, the interval of the minimum width containing 97.5% of the area of the marginal distributions. We ran the MCWLA using all the 55,000 sets of parameters sampled by the final run of the M–H algorithm to investigate the uncertainties of the ensemble prediction. From the 55,000 sets of parameters, we further selected the optimal parameter set that produces the minimum root mean-square error (RMSE) between modelled and observed historical crop-yield series.

4. Results

4.1. *Inversion results of model parameters and the optimal parameter set*

Our inversion results of model parameters at the grid of Harbin for spring maize in the Northeast China Plain and at the grid of Zhengzhou for summer maize in the North China Plain are shown in Table 1. We list the model parameters' 97.5% high-probability intervals, mean estimates, standard deviations, and the optimal parameter set, based on the 55,000 sets of parameters sampled by the final run of the M–H algorithm. Some other parameters or constants used in the study are listed in Table 2. These parameters are used for model evaluation and uncertainties analysis.

4.2. *Model evaluation*

First, using the corresponding optimal parameter set, the MCWLA was run at each $0.5^\circ \times 0.5^\circ$ grid with maize cultivation fraction $\ge$0.05 across the two major production provinces for

Table 2 – Values of some model parameters or constants used in the study.

Symbol	Description	Values used in the study	References
W_{sow}	Threshold fraction of soil water for automatic sowing	0.5	This study
L_{AImax}	Maximum leaf area index	5.8 $m^2\ m^{-2}$	Cavero et al. (2000)
F_{cr}	A parameter to adjust the damage extent of one flooding event	5.0	This study
L_{AIdg}	Mean rate of L_{AI} decrease after flowering	−0.002 $m^2\ m^{-2}$	This study
D_{rmax}	Maximum root depth	1.5 m	Cavero et al. (2000)
ξ	Priestley–Taylor coefficient	1.32	Priestley and Taylor (1972)
I_c	Interception storage parameter	0.01	Kergoat (1998)
c_p	Specific heat of moist air	1.013 kJ $kg^{-1}\ °C^{-1}$	Allen et al. (1998b)
p_{re}	Atmospheric pressure	100 kPa	Sellers et al. (1996)
ε	Molecular weight ratio of water vapour/dry air	0.622	Allen et al. (1998b)
ρ	Density of air	1.225 kg m^{-3}	Sellers et al. (1996)
a	Leaf respiration as a fraction of Rubisco capacity	For C_3 plants 0.015, for C_4 plants 0.02	Farquhar et al. (1980)
k_c	Michaelis constant for CO_2 at 25 °C	30 Pa (Q_{10} = 2.1)	Collatz et al. (1991)
k_o	Michaelis constant for O_2 at 25 °C	30 kPa (Q_{10} = 1.2)	Collatz et al. (1991)
τ	CO_2/O_2 specificity ratio at 25 °C	2600 (Q_{10} = 0.57)	Brooks and Farquhar (1985)
m_c	Moisture content of grain	0.11	NRC (1982)
c_c	Carbon content of biomass	0.45	Schlesinger (1997)
g_{min}	Minimum canopy conductance	0.5 mm s^{-1}	Haxeltine and Prentice (1996b)
k_b	Light extinction coefficient	0.5	Woodward (1987)
θ	Co-limitation parameter	0.7	McMurtrie and Wang (1993)
α_m	Empirical parameter in calculating E_{demand}	1.391	Monteith (1995)

spring maize (i.e., Heilongjiang and Jilin provinces) and the two major production provinces for summer maize (i.e., Henan, and Shandong provinces), respectively, resulting in a deterministic yield prediction (YdOp) for each grid. Then the MCWLA was run using all the 55,000 sets of parameters sampled by the M–H algorithm, and an ensemble mean yield prediction (YdEn) for each grid can be derived by averaging the output from each set of parameters. We calculated the modelled sowing-area-weighted yield for each province using the modelled yields and maize growing area ratio (to province total) at the grids (Qiu et al., 2003) across the province, assuming the yearly growing area ratio at each grid (to province total) did not change throughout the period. The performance of the model was evaluated by calculating the Pearson correlation coefficient (r) and RMSE between the modelled (YdOp or YdEn) and the corresponding observed yield series at both the crop model grid scale and province scale. Correlations are considered to be significant at $p < 0.05$.

Table 3 – The r and RMSE (kg ha^{-1}) between the modelled (YdOP and YdEn) and observed yield series at some crop model grids and at province scale (in *italic*).

Province/grid	YdOp r	YdEn r	YdOp RMSE	YdEn RMSE	Years
Heilongjiang province	0.68**	0.67**	388	419	1985–2002
Harbin	0.74	0.61	712	933	1997–2002
Mudanjiang	0.10	0.42	954	883	1995–2002
Jilin	0.45	0.52*	859	951	1985–2002
Yanji	0.53	0.54	1845	1529	1992–2002
Changchun	0.18	0.41	1536	1544	1995–2002
Tonghua	0.67	0.64	800	781	1996–2002
Siping	0.40	0.88**	1146	1391	1996–2002
Henan province	0.48	0.57*	501	563	1987–2002
Luoyang	0.03	0.17	1389	1334	1987–2002
Pingdingshan	0.38	0.41	841	882	1992–2002
Luohe	0.30	0.13	1322	1501	1994–2002
Xinxiang	0.33	0.18	820	1072	1994–2002
Shandong province	0.59*	0.82**	439	309	1985–2002
Jinan	0.52	0.62*	684	756	1989–2002
Qingdao	0.47	0.61*	1329	1225	1991–2002
Weifang	0.30	0.51	634	648	1992–2002
Taian	0.17	0.33	1350	1283	1993–2002

* $p < 0.05$.
** $p < 0.01$.

4.2.1. *Model skill at the grid scale*

For spring maize, the observed data at the grid of Harbin from 1985 to 1996 were used for model calibration. In contrast, the observed data at the grid of Harbin from 1997 to 2002; the grid of Mudanjiang from 1995 to 2002; and the grids of Yanji from 1992 to 2002, Changchun from 1995 to 2002, Tonghua from 1996 to 2002 and Siping from 1996 to 2002 were used for model evaluation (Table 3). At the grid of Harbin, the r between the modelled and observed yield series is 0.61 and 0.74 for YdEn and YdOp, respectively; the RMSE is 933 and 712 kg ha^{-1} for YdEn and YdOp, respectively (Fig. 2a). At the grid of Mudanjiang, the r between the modelled and observed yield series is 0.42 and 0.10 for YdEn and YdOp, respectively; the RMSE is 883 and 954 kg ha^{-1} for YdEn and YdOp, respectively (Fig. 2b). The r and RMSE between the modelled and observed yield series at all the selected grids including Yanji, Changchun, Tonghua and Siping are listed in Table 3. Agreement between observed and modelled yield was variable, with r ranging from 0.41 to 0.88 ($p < 0.01$) for YdEn and from 0.10 to 0.74 for YdOp.

For summer maize, the observed data at the grid of Zhengzhou from 1995 to 2002 were used for model calibration. In contrast, the observed data at the grid of Luoyang from 1987 to 2002; at the grid of Pingdingshan from 1992 to 2002; and at the grids of Luohe from 1994 to 2002, Xinxiang from 1994 to 2002, Jinan from 1989 to 2002, Qingdao from 1991 to 2002,

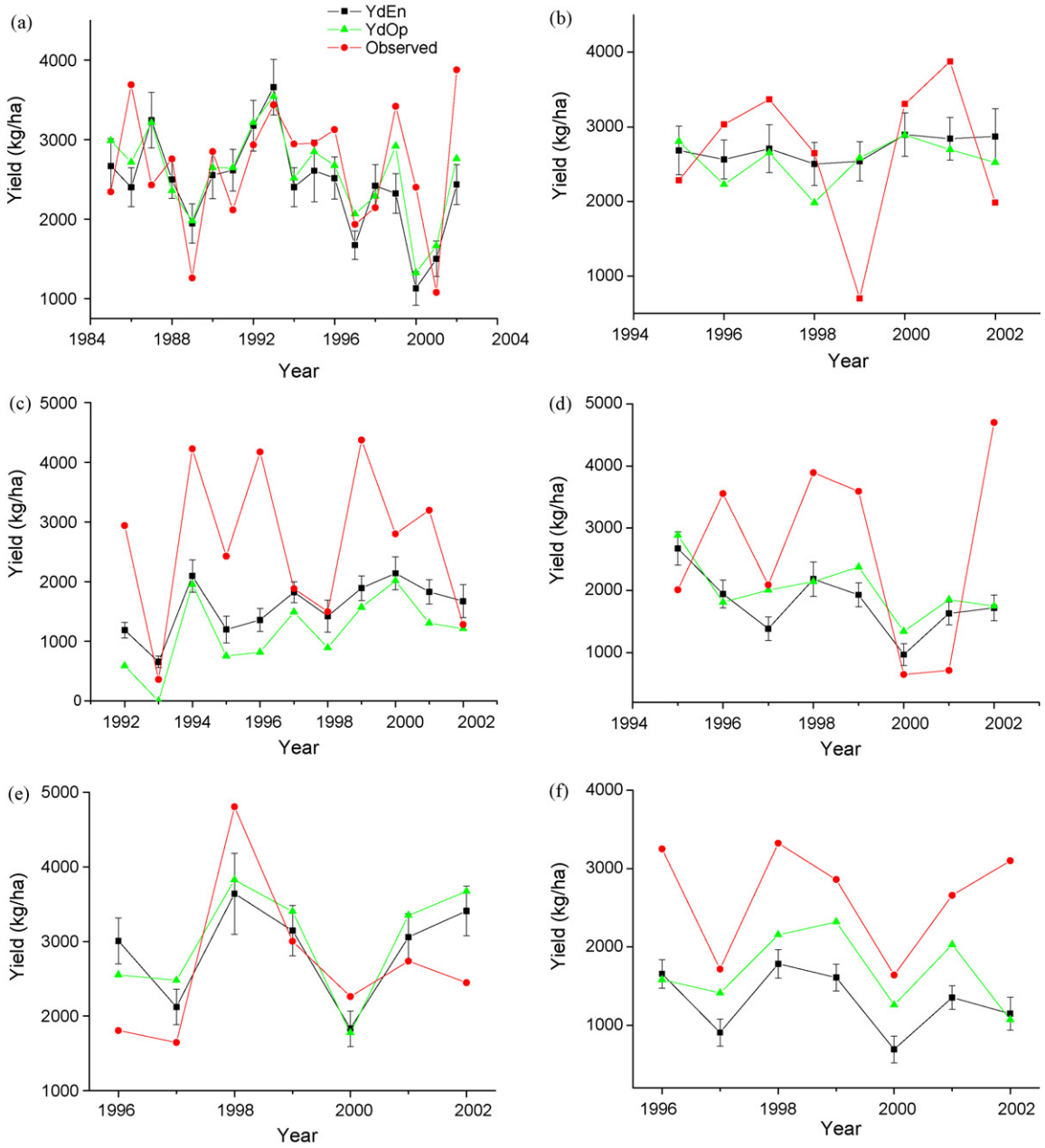


Fig. 2 – Time series in the modelled and observed yield at the crop model grid scale for spring maize at the grid of Harbin (a), Mudanjiang (b), Yanji (c), Changchun (d), Tonghua (e), and Siping (f). YdEn, ensemble yield prediction; YdOp, deterministic yield prediction.

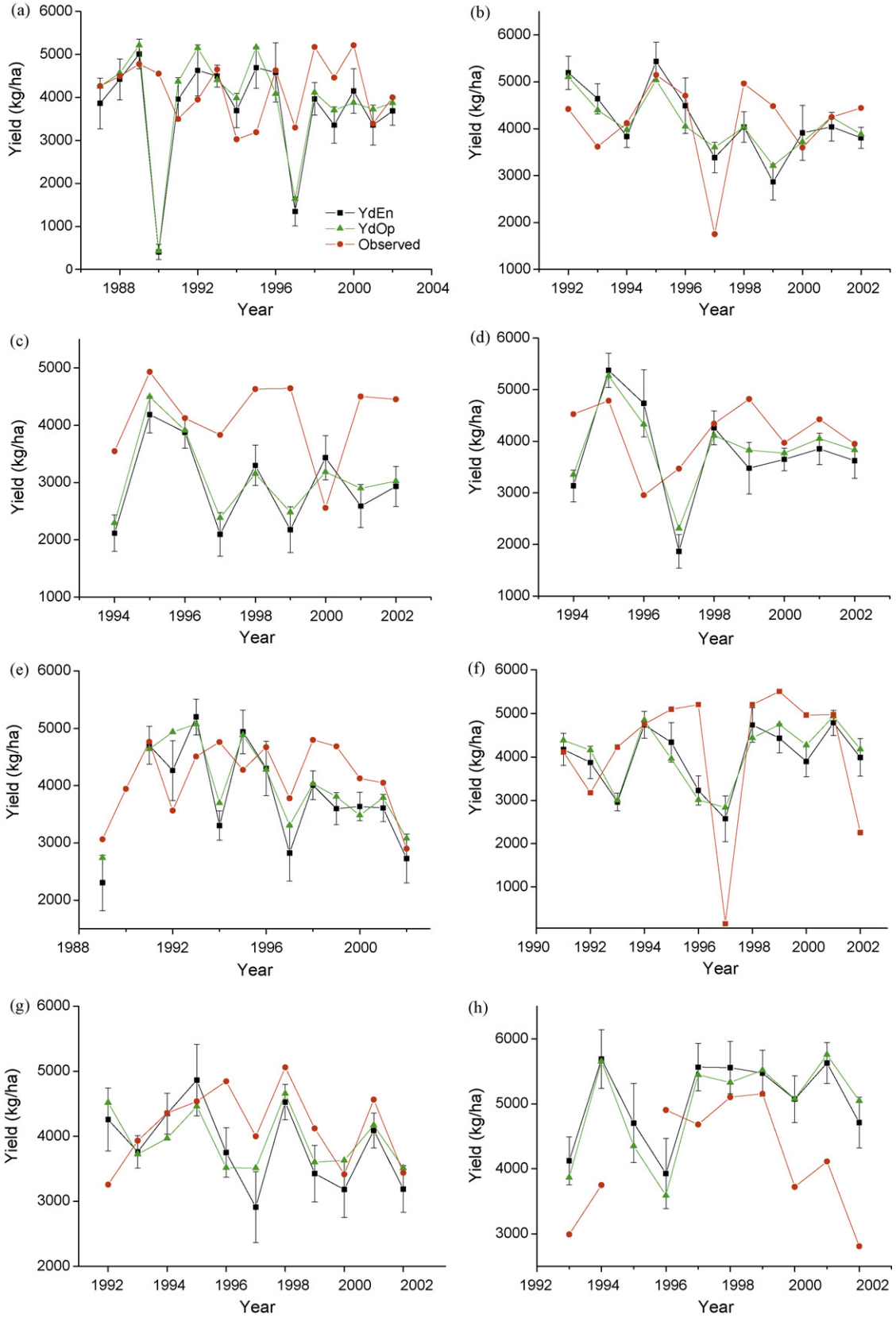


Fig. 3 – Time series in the modelled and observed yield at the crop model grid scale for summer maize at the grid of Luoyang (a), Pingdingshan (b), Luohe (c), Xinxiang (d), Jinan (e), Qingdao (f), Weifang (g), and Taian (h). YdEn, ensemble yield prediction; YdOp, deterministic yield prediction.

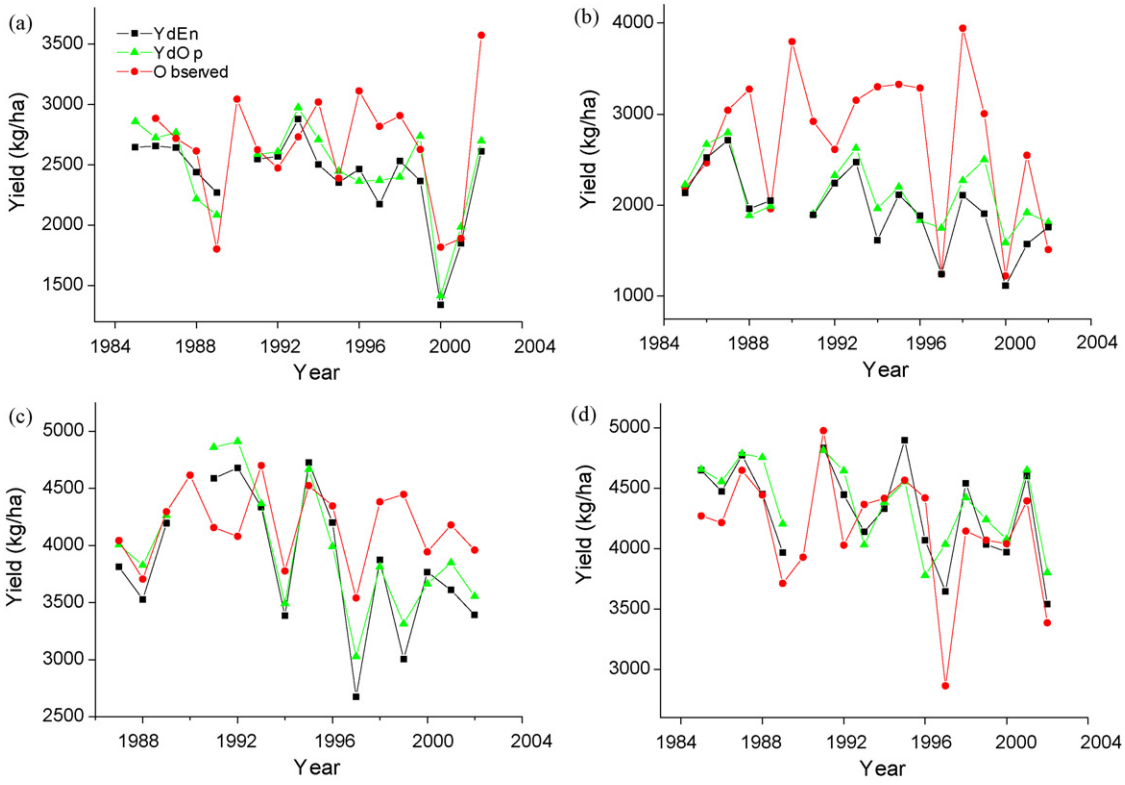


Fig. 4 – Time series in the modelled and observed yield at the province scale for Heilongjiang province (a), Jilin province (b), Henan province (c), and Shandong province (d). YdEn, ensemble yield prediction; YdOp, deterministic yield prediction.

Weifang from 1992 to 2002 and Taian from 1993 to 2002 were used for model evaluation. At the grid of Luoyang, the r between the modelled and observed yield series is 0.17 and 0.03 for YdEn and YdOp, respectively; the RMSE is 1334 and 1389 kg ha^{-1} for YdEn and YdOp, respectively (Fig. 3a). At the grid of Pingdingshan, the r between the modelled and observed yield series is 0.41 and 0.38 for YdEn and YdOp, respectively; the RMSE is 882 and 841 kg ha^{-1} for YdEn and YdOp, respectively (Fig. 3b). The r and RMSE between the modelled and observed yield series at all the selected grids including Luohe, Xinxiang, Jinan, Qingdao, Weifang and Taian are also listed in Table 3. The r ranged from 0.13 to 0.62 ($p < 0.05$) for YdEn and from 0.03 to 0.52 for YdOp. The r was significant at the 0.05 level at several grids in Shandong province. The RMSE can be further minimized by bias correction based on observations, although r cannot.

4.2.2. *Model skill at the province scale*

The performance of the model was further evaluated at the province scale. For spring maize in Heilongjiang province from 1985 to 2002, the r between the modelled and observed yield series is 0.67 ($p < 0.01$) and 0.68 ($p < 0.01$) for YdEn and YdOp, respectively; the RMSE is 419 and 388 kg ha^{-1} for YdEn and YdOp, respectively (Table 3) (Fig. 4a). In Jilin province from 1985 to 2002, the r between the modelled and observed yield series is 0.52 ($p < 0.05$) and 0.45 for YdEn and YdOp, respectively; the RMSE is 951 and 859 kg ha^{-1} for YdEn and YdOp, respectively (Fig. 4b).

For summer maize in Henan province from 1987 to 2002, the r between the modelled and observed yield series is 0.57 ($p < 0.05$) and 0.48 for YdEn and YdOp, respectively; the RMSE is 563 and 501 kg ha^{-1} for YdEn and YdOp, respectively (Table 3) (Fig. 4c). In Shandong province from 1985 to 2002, the r between the modelled and observed yield series is 0.82 ($p < 0.01$) and 0.59 ($p < 0.05$) for YdEn and YdOp, respectively; the RMSE is 309 and 439 kg ha^{-1} for YdEn and YdOp, respectively (Fig. 4d).

The ensemble hindcasts by MCWLA captured significantly the interannual variability of maize yield in all the four province from 1985 to 2002 (Table 3). Among other things, the relative performance of the MCWLA within an individual province could be attributed to the relative crop irrigation fraction, because the present version of the MCWLA does not account for irrigation. For example, the maize irrigation fraction in Henan province (≈0.5) is quite higher than that in Heilongjiang province (<0.2), which led to a relatively bad performance of the MCWLA in Henan province (Fig. 4c).

5. Discussion

5.1. *Crop response to elevated [CO_2]*

Extensive controlled-environment experiments have showed that elevated [CO_2] lead to a decrease in stomatal conductance in both C_3 and C_4 species (Rogers et al., 1983; Morrison and Gifford, 1984a,b; Morrison, 1987; Bunce, 1996), which reduces

the transpiration rate per unit leaf area. Morrison and Gifford (1984b) found that stomatal conductance was reduced over a range of species by 36% while transpiration was reduced by 21%, the difference being attributed to the higher leaf temperatures. Similar average values of 34% and 23% for stomatal conductance and transpiration were found in a literature survey by Cure and Acock (1986). Both an increase in photosynthesis and a decrease in transpiration result in an increase in a plant's water use efficiency, the ratio of carbon fixation to water loss. A review of 18 crop species in controlled environments (Kimball and Idso, 1983) suggested that water use efficiency might double with the doubling of CO_2. The enhancement in CO_2 effects on growth and water use efficiency when soils dry results partly from slower transpiration and a delay in the onset of drought (Allen et al., 1998a). This is especially true of C_4 species, many of which exhibit little photosynthetic response to CO_2 until soil begins to dry (Gifford and Morison, 1985). Leaf area of maize did not respond to CO_2 when well-watered, but increase by up to 35% at elevated [CO_2] as soil dried (Samarakoon and Gifford, 1996). Plant biomass responded similarly (Samarakoon and Gifford, 1996). Leakey et al. (2004) showed maize growth at elevated [CO_2] significantly increased leaf photosynthetic CO_2 uptake rate by up to 41%, and 10% on average. Kim et al. (2006, 2007) also showed that CO_2 enrichment (from 370 to 750 ppm) did not enhance the growth (including leaf area per plant, specific leaf area, biomass and its allocation) or canopy photosynthesis of maize plants, however leaves grown at elevated [CO_2] exhibited over 50% reduction in stomatal conductance and transpiration, and canopy evapotranspiration rates decreased by 22% from emergence to silking. Water use efficiency increased by 108%.

The MCWLA captures the key responses mechanism quite well (Figs. 5 and 6). When atmospheric [CO_2] changed from 370 to 750 ppm, for spring maize at the grid of Harbin (Fig. 5), g_c and T_T reduced by 26.6% (18.5%) and 44.5% (38.1%) on average, respectively, during the growing period in 2002 with total precipitation 476 mm (1997 with total precipitation 490 mm); L_{AI} and crop yield increased by 0.96% (5.56%) and 3.25% (6.15%) on average, respectively, in 2002 (1997). For summer maize at the grid of Zhengzhou (Fig. 6), g_c and T_T reduced by 31.0% (26.6%) and 50.7% (49.4%) on average, respectively, during the growing period in 2002 with total precipitation 701 mm ((until flowering in 1997 with total precipitation 354 mm); L_{AI} and crop yield increased by 0.0% (0.33%) and 0.0% (24.25%) on average, respectively, in 2002 (1997). The results suggest water use efficiency increased by 86.0% (71.5%) on average at Harbin and by 102.8% (145.6%) on average at Zhengzhou in 2002 (1997). A delay in the onset of drought by elevated [CO_2] also was simulated at Zhengzhou in 1997 (Fig. 6).

5.2. V_{PD}, T_T, and crop yield

V_{PD} is another important variable that affects T_T and consequently water use and crop yield (Challinor and Wheeler, 2008a). The MCWLA simulates the relationship between

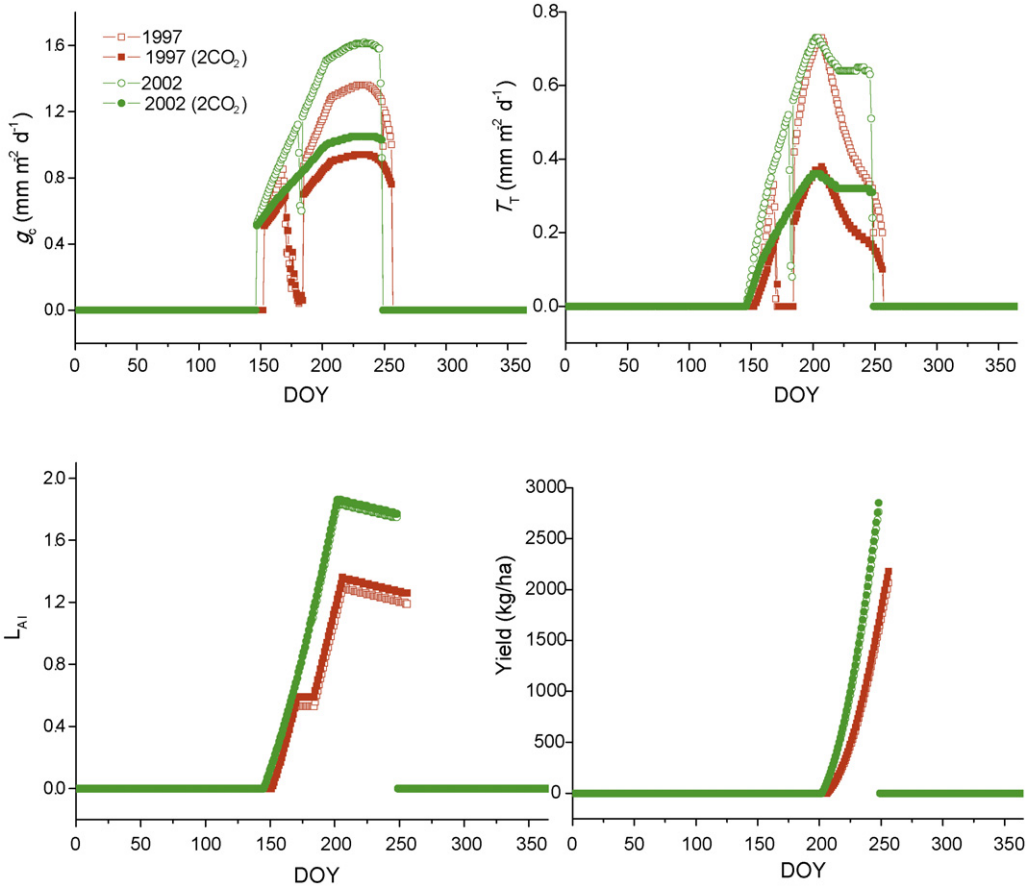


Fig. 5 – MCWLA simulated daily changes in g_c, T_T, L_{AI} and crop yield of spring maize at baseline [CO_2] (370 ppm) and elevated [CO_2] (750 ppm) at the grid of Harbin in 1997 and 2002.

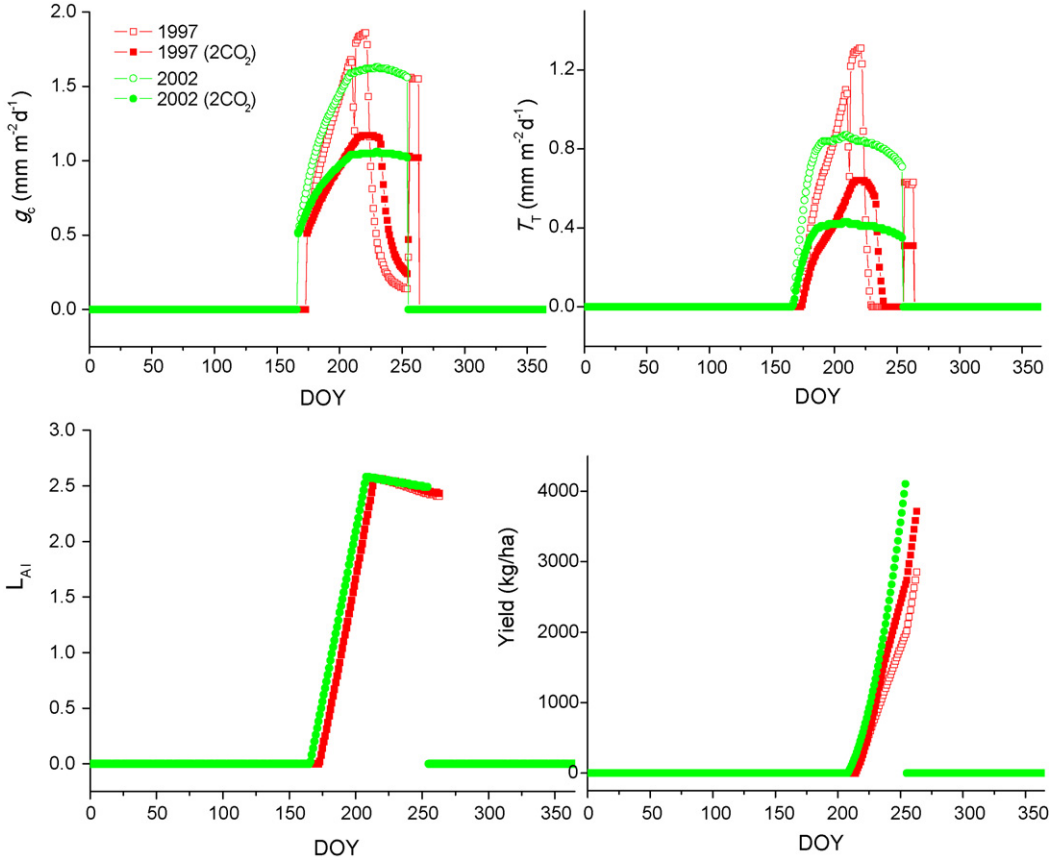


Fig. 6 – As for Fig. 5 but summer maize at the grid of Zhengzhou.

V_{PD} and T_T using Eqs. (34) and (35), which also includes indirectly the effects of soil moisture through g_c. Crop T_T increase with increasing V_{PD}, however the increase has limits and a limiting maximum T_T is commonly reached at a V_{PD} of ~2.0 kPa. (McNaughton and Jarvis, 1991; Fletcher et al., 2007). Bunce (1981) showed decreased g_c and T_T in a number of species between 1.0 and 2.5 kPa. Although these studies showed a similar pattern, the T_T response differs both among and within species (Isoda and Wang, 2002). The MCWLA also captures the T_T response quite well under both atmospheric [CO_2] (370 and 750 ppm) (Fig. 7). At the grid of Harbin, T_T increased with V_{PD} and reached a maximum T_T at a V_{PD} of about 0.95 kPa in 1997 and about 0.75 kPa in 2002, then decreased with V_{PD} increasing (Fig. 7). At the grid of Zhengzhou, T_T increased with V_{PD} and reached a maximum T_T at a V_{PD} of about 0.98 kPa in 1997, and about 0.87 kPa in 2002, then decreased with V_{PD} increasing (Fig. 7). Soil drought could complicate the response pattern, as in 1997 (Fig. 7a and c).

5.3. *Uncertainties in model parameters and yield prediction*

Uncertainties in model parameters are presented in Table 1. As a result, ensemble predictions (by perturbing the parameters) produce a large yield range, for example, with standard errors ranging from 179 to 390 kg ha^{-1} in Harbin and from 178 to 634 kg ha^{-1} in Luoyang (Fig. 2). In this study, the model parameters were calibrated at the representative grid cells (Harbin and Zhengzhou) for spring maize and summer maize, and then applied in the nearby two provinces, respectively. Ensemble predictions allow for accounting the physical and biological uncertainty (Challinor et al., 2005a). The optimal parameter set worked better at some grids and provinces; in contrast, ensemble predictions work better at other grids and provinces, suggesting the optimal parameter set was locally specific. At province scale, ensemble hindcasts captured significantly the yield variability in all the four investigated provinces. Ideally the model parameters PDF and the optimal parameter set are calibrated against the historical datasets at the same grid or a large area before the model is used for predictions in the target grid or a large area. In addition, there are many other nonclimatic factors affecting the weather-yield correlations (Challinor et al., 2005b), such as changes in the fraction of the crop under irrigation or in cultivar-specific properties. Although the statistical data on crop growing area and yield are the best source for large-area studies, the accuracy of the data may have measurable uncertainties and may change over time (Challinor et al., 2005b).

Because uncertainties in model parameters affect assessments of the impact of climate variability, the Bayesian probability inversion and an MCMC technique is an effective method to analyze the uncertainties in parameter estimation and model prediction. Along this line, we plan to further develop a super-ensemble-based probabilistic projection to

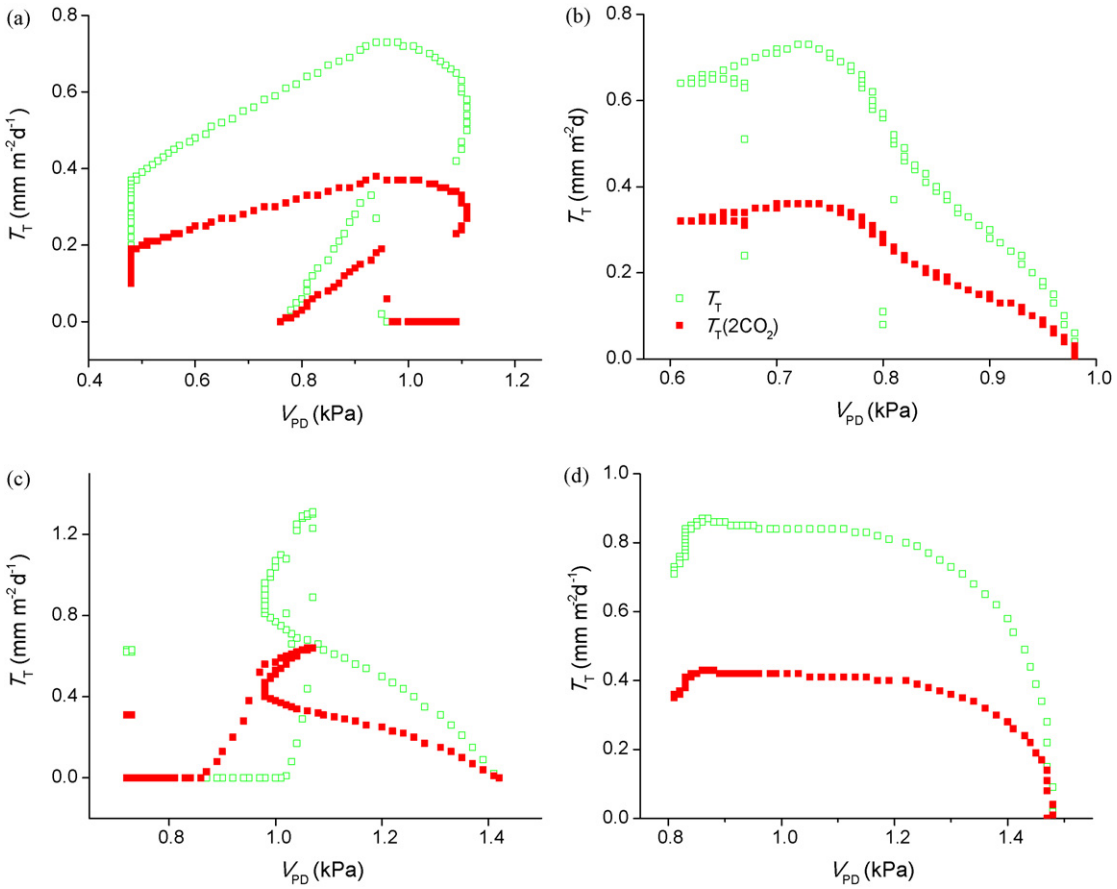


Fig. 7 – The relationship between V_{PD}, and T_T at baseline [CO_2] (370 ppm) and elevated [CO_2] (750 ppm) simulated by the MCWLA at the grid of Harbin in 1997(a) and 2002 (b), and at the grid of Zhengzhou in 1997 (c) and 2002 (d).

account for the uncertainties not only from the climate and emission scenarios (Tao et al., 2008b), but also from the biophysical parameters.

5.4. *Climate variability and crop production prediction over a large area*

The MCWLA was developed to examine the impacts of climate variability on crop phenology and yield over a large area. Among the key impact mechanisms of climate change, the MCWLA accounts mechanically for the impacts of climate variables and elevated [CO_2] on canopy net photosynthesis, stomatal conductance and T_T, instead of using proportionality factors as do many crop models (Long et al., 2006; Tubiello et al., 2007b).

The MCWLA also captures the impacts of mean temperature on crop phenology change. Although the present version of the MCWLA does not explicitly simulate the high temperature stress on crop yield, as did Horie et al. (1995) and Challinor et al. (2005c), it does account for the impacts of extreme temperature stress on photosynthesis and subsequently on stomatal conductance, transpiration, and crop yield.

Level of complexity in crop modelling is closely related to the focus and purpose of the model. Complexity is not a prerequisite for quantifying the impacts of elevated [CO_2] and its interaction with water stress (Tubiello and Ewert, 2002; Challinor and Wheeler, 2008b), however the models that include the processes and interactions that are significant determinant of crop water use and yield could be important, especially for future climate. The MCWLA simulates the changes of water use efficiency with climate and [CO_2] intrinsically and consequently is internally consistent. In contrast, GLAM (Challinor et al., 2005a; Challinor and Wheeler, 2008b) simulates the effects of climate change and elevated [CO_2] in a manner of 'offline' by adopting a new parameter set. The robust, process-based representation of the coupled CO_2 and H_2O exchanges used in the MCWLA have been validated over the large scale including agriculture ecosystem (Haxeltine and Prentice, 1996a,b; Sitch et al., 2003; Bondeau et al., 2007). Many parameters in MCWLA as listed in Table 2 can be applied universally or with small changes. The MCWLA also simplifies the modelling of the impacts due to factors other than weather using a single yield-gap parameter, as in GLAM (Challinor et al., 2004). All of these make the MCWLA suitable for examining the impacts of climate variability on crop phenology and yield over a large area both in present and future climate condition.

6. Conclusions

A new process-based crop model, the MCWLA, was developed to capture crop–weather relationships over a large area. Because the MCWLA includes robust process-based represen-

tation of the coupled CO_2 and H_2O exchanges, it can capture mechanically the impacts of V_{PD}, soil moisture, temperature and elevated [CO_2] on canopy net photosynthesis, stomatal conductance, and transpiration, which is crucial for the models that account for the impacts of [CO_2] and drought on water use and crop production. Ensemble hindcasts (by perturbing parameters) and deterministic hindcasts (using the optimal parameters set) showed that the MCWLA could capture the interannual variability of crop yield quite well, especially at a large scale. Furthermore, MCWLA's simulations on crop response to elevated [CO_2] agree well with the controlled-environment experiments, suggesting its validity in future climate.

MCWLA simplifies the modelling of the impacts due to factors other than weather. Also many parameters in MCWLA can be applied universally or with small changes. Therefore, MCWLA can be easily extended to other crop and/or regions to examine the impacts of climate variability on crop phenology and yield over a large area both in present and future climate condition.

The Bayesian probability inversion and an MCMC technique were applied to the MCWLA to analyze uncertainties in parameter estimation and model prediction and to optimize the model. We demonstrated that the system is effective in developing an ensemble-based probabilistic projection, an optimal projection, and to analyze the uncertainties.

Acknowledgements

This study was supported by National Key Programme for Developing Basic Science (Project Number 2009CB421105), China and the Innovative Program of Climate Change Projection for the 21st Century (KAKUSHIN Program), Japan. F. Tao acknowledges the support of the ''Hundred Talents'' Program of the Chinese Academy of Sciences. We appreciate Dr. T. Iizumi of the National Institute for Agro-Environmental Sciences for sharing the MCMC programme. We are grateful to the three anonymous reviewers and editor for their insightful comments on an earlier version of this manuscript.

REFERENCES

Ainsworth, E.A., Davey, P.A., Bernacchi, C.J., Dermody, O.C., Heaton, E.A., Moore, D.J., Morgan, P.B., Naidu, S.L., Yoo Ra, H., Zhu, X., Curtis, P.S., Long, S.P., 2002. A meta-analysis of elevated [CO_2] effects on soybean (*Glycine max*) physiology, growth and yield. Global Change Biol. 8, 695–709.

Allen Jr., L.H.R.R., Jones, J.W., Jones, P.H., 1998a. Soybean leaf water potential responses to carbon dioxide and drought. Agron. J. 90, 375–383.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998b. Crop Evapotranspiration—Guideline for Computing Crop Water Requirements-FAO Irrigation and Drainage Paper 56. FAO, Rome, Italy.

Atkinson, C.J., Wookey, P.A., Mansfield, T.A., 1991. Atmospheric pollution and the sensitivity of stomata on barley leaves to abscissic acid and carbon dioxide. New Phytol. 117, 535–541.

Bannayan, M., Kobayashib, K., Kim, H., Lieffering, M., Okada, M., Miura, S., 2005. Modeling the interactive effects of atmospheric CO_2 and N on rice growth and yield. Field Crops Res. 93, 237–251.

Bondeau, A., Smith, P.C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., Lotze-Campen, H., Müller, C., Reichstein, M., Smith, B., 2007. Modelling the role of agriculture for the 20th century global terrestrial carbon balance. Global Change Biol. 13, 679–706.

Braswell, B.H., Sacks, W.J., Linder, E., Schimel, D.S., 2005. Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations. Global Change Biol. 11, 335–355.

Brooks, A., Farquhar, G.D., 1985. Effects of temperature on the CO_2/O_2 specificity of ribulose-1,5 bisphosphate carboxylase/oxygenase and the rate of respiration in the light. Planta 165, 397–406.

Bunce, J.A., 1981. Comparative responses of leaf conductance to humidity in single attached leaves. J. Exp. Bot. 32, 629–634.

Bunce, J.A., 1996. Growth at elevated carbon dioxide concentration reduces hydraulic conductance in alfalfa and soybean. Global Change Biol. 2, 155–158.

Cavero, J., Farre, I., Debaeke, P., Faci, J.M., 2000. Simulation of maize yield under water stress with EPIC phase and CROPWAT models. Agron. J. 92, 679–690.

Challinor, A.J., Wheeler, T.R., 2008a. Crop yield reduction in the tropics under climate change: processes and uncertainties. Agric. For. Meteorol. 148, 343–356.

Challinor, A.J., Wheeler, T.R., 2008b. Use of a crop model ensemble to quantify CO_2 stimulation of water-stressed and well-watered crops. Agric. For. Meteorol. 148, 1062–1077.

Challinor, A.J., Wheeler, T.R., Craufurd, P.Q., Slingo, J.M., Grimes, D.I.F., 2004. Design and optimisation of a large-area process-based model for annual crops. Agric. For. Meteorol. 124, 99–120.

Challinor, A.J., Wheeler, T.R., Slingo, J.M., Hemming, D., 2005a. Quantification of physical and biological uncertainty in the simulation of the yield of a tropical crop using present-day and doubled CO_2 climates. Phil. Trans. R. Soc. B 360, 2085–2094.

Challinor, A.J., Wheeler, T.R., Slingo, J.M., Craufurd, P.Q., Grimes, D.I.F., 2005b. Simulation of crop yield using EAR-40: limits to skill and nonstationarity in weather-yield relationships. J. Appl. Meteorol. 44, 516–531.

Challinor, A.J., Wheeler, T.R., Craufurd, P.Q., Slingo, J.M., 2005c. Simulation of the impact of high temperature stress on annual crop yields. Agric. For. Meteorol. 135, 180–189.

Collatz, G.J., Ball, J.T., Grivet, C., Berry, A., 1991. Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer. Agric. For. Meteorol. 54, 107–136.

Collatz, G.J., Ribas-Carbo, J., Berry, J.A., 1992. Coupled photosynthesis–stomatal conductance model for leaves of C4 plants. Aust. J. Plant Physiol. 19, 519–538.

Cramer, W., Bondeau, A., Woodward, F.I., Prentice, I.C., Betts, R.A., Brovkin, V., Cox, P.M., Fisher, V., Foley, J.A., Friend, A.D., Kucharik, C., Lomas, M.R., Ramankutty, N., Sitch, S., Smith, B., White, A., Young-Molling, C., 2001. Global response of terrestrial ecosystem structure and function to CO_2 and climate change: results from six dynamic global vegetation models. Global Change Biol. 7, 357–373.

Cure, J.D., Acock, B., 1986. Crop responses to carbon dioxide doubling: a literature survey. Agric. For. Meteorol. 38, 127–145.

Dewar, R.C., 1996. The correlation between plant growth intercepted radiation: an interpretation in terms of optimal plant nitrogen content. Ann. Bot. 78, 125–136.

Doorenbos, J., Kassam, A.H., 1979. Yield Response to Water. FAO Irrigation and Drainage 33. FAO, Viale delle Terme di Caracalla, Rome, Italy.

Farquhar, G.D., von Caemmerer, S., 1982. Modelling of photosynthetic response to environmental conditions. In: Lange, O.L., Nobel, P.S., Osmond, C.B., Ziegler, H. (Eds.), Encyclopedia of Plant Physiology, vol. 12B. Springer-Verlag, Berlin, pp. 549–587.

Farquhar, G.D., von Caemmerer, S., Berry, J.A., 1980. A biochemical model of photosynthetic CO_2 assimilation in leaves of C3 species. Planta 149, 78–90.

Fischer, G., Shah, M., van Velthuizen, H., 2002. Climate change and agricultural vulnerability. Technical Report of the International Institute for Applied Systems Analysis. Available at http://www.iiasa.ac.at/Research/LUC/.

Fletcher, A.L., Sinclair, T.R., Hartwell Allen Jr., L., 2007. Transpiration responses to vapor pressure deficit in well watered 'slow-wilting' and commercial soybean. Environ. Exp. Bot. 61, 145–151.

Food and Agriculture Organization of the United Nations (FAO), 1991. The Digitized Soil Map of the World (Release 1.0), vol. 67/1. FAO, Rome.

Gale, M.R., Grigal, D.F., 1987. Vertical root distributions of northern tree species in relation to successional status. Can. J. For. Res. 17, 829–834.

Gelfand, A.E., Smith, A.F.M., 1990. Sampling-based approaches to calculating marginal densities. J. Am. Stat. Assoc. 85, 398–409.

Geman, S., Geman, D., 1984. Stochastic relaxation Gibbs distributions and the Bayesian restoration of images. IEEE Trans. Pattern Anal. Mach. Intel. 6, 721–741.

Gifford, R.M., Morison, J.I.L., 1985. Photosynthesis, water use and growth of a C_4 grass stand at high CO_2 concentration. Photosynth. Res. 7, 69–76.

Gill, J., 2002. Bayesian Methods: A Social and Behavioral Approach. CRC Press, Boca Raton, FL.

Hansen, J.W., Jones, J.W., 2000. Scaling-up crop models for climatic variability applications. Agric. Syst. 65, 43–72.

Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chain and their applications. Biometrika 57, 97–109.

Haxeltine, A., Prentice, I.C., 1996a. BIOME3: an equilibrium terrestrial biosphere model based on ecophysiological constraints, resources availability, and competition among plant functional types. Global Biogeo. Cycles 10, 693–709.

Haxeltine, A., Prentice, I.C., 1996b. A general model for the light-use efficiency of primary production. Funct. Ecol. 10, 551–561.

Horie, T., Nakagawa, H., Centeno, H.G.S., Kropff, M.J., 1995. The rice crop simulation model SIMRIW and its testing. In: Matthews, R.B., Kropff, M.J., Bachelet, D. (Eds.), Modelling the Impact of Climate Change on Rice Production in Asia. CAB International, Oxon, UK, pp. 51–66.

Hunt, E.R., 1994. Relationship between woody biomass and PAR conversion efficiency for estimating net primary production from NDVI. Int. J. Remote Sen. 15, 1725–1730.

Intergovernmental Panel on Climate Change (IPCC), 2001. Climate Change 2001: Impacts, Adaptation, and Vulnerability. Cambridge University Press, Cambridge, UK, pp. 556–560.

Isoda, A., Wang, P., 2002. Leaf temperature and transpiration of field grown cotton and soybean under arid and humid conditions. Plant Prod. Sci. 5, 224–228.

Iizumi, T., Yokozawa, M., Nishimori, M. Parameter estimation and uncertainty analysis of a large-scale crop model for paddy rice: application of a Bayesian approach. Agric. For. Meteorol., in press.

Jagtap, S.S., Jones, J.W., 2002. Adaptation and evaluation of the CROPGRO-soybean model to predict regional yield and production. Agric. Ecosyst. Environ. 93, 73–85.

Jamieson, P.D., Semenov, M.A., 2000. Modelling nitrogen uptake and redistribution in wheat. Field Crops Res. 68, 21–29.

Jarvis, P.G., MacNaughton, K.G., 1986. Stomata control of transpiration: scaling up from leaf to region. Adv. Ecol. Res. 15, 1–49.

Jarvis, A.J., William, J.D., 1998. The coupled response of stomatal conductance to photosynthesis and transpiration. J. Exp. Bot. 49, 399–406.

Kergoat, L., 1998. A model of hydrologic equilibrium of leaf area index on a global scale. J. Hydrol. 212–213, 267–286.

Kim, S.H., Sicher, R.C., Bae, H., Gitz, D.C., Baker, J.T., Timlin, D.J., Reddy, V.R., 2006. Canopy photosynthesis, evapotranspiration, leaf nitrogen, and transcription profiles of maize in response to CO_2 enrichment. Global Change Biol. 12, 588–600.

Kim, S.H., Gitz, D.C., Sicher, R.C., Baker, J.T., Timlin, D.J., Reddy, V.R., 2007. Temperature dependence of growth, development, and photosynthesis in maize under elevated CO_2. Environ. Exp. Bot. 61, 224–236.

Kimball, B.A., Idso, S.B., 1983. Increasing atmospheric CO_2: effects on crop yield, water use and climate. Agric. Water Manag. 7, 55–72.

Kimball, B.A., Pinter Jr., P.J., Garcia, R.L., LaMorte, R.L., Wall, G.W., Hunsaker, D.J., Wechsung, G., Wechsung, F., Kartschall, T., 1995. Productivity and water use of wheat under free-air CO_2 enrichment. Global Change Biol. 1, 429–442.

Knorr, W., Kattge, J., 2005. Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling. Global Change Biol. 11, 1333–1351.

Larcher, 1983. Physiological Plant Ecology. Springer, Heidelberg.

Leakey, A.D.B., Bernacchi, C.J., Dohleman, F.G., Ort, D.R., Long, S.P., 2004. Will photosynthesis of maize (*Zea mays*) in the US Corn Belt increase in future [CO_2] rich atmospheres? An analysis of diurnal courses of CO_2 uptake under free-air concentration enrichment (FACE). Global Change Biol. 10, 951–962.

Leonard, T., Hsu, J.S.J., 1999. Bayesian Methods: An Analysis for Statistics and Interdisciplinary Researchers. Cambridge University Press, New York.

Li, K.Y., Jong, R.D., Coe, M.T., Ramankutty, N., 2006. Root-water-uptake based upon a new water stress reduction and an asymptotic root distribution function. Earth Interactions, Paper 10-014.

Lobell, D.B., Hicke, J.A., Asner, G.P., Field, C.B., Tucker, C.J., Los, S.O., 2002. Satellite estimates of productivity and light use efficiency in United States agriculture 1982–1998. Global Change Biol. 8, 722–735.

Long, S.P., Ainsworth, E.A., Rogers, A., Ort, D.R., 2004. Rising atmospheric carbon dioxide: plants FACE the future. Annu. Rev. Plant Biol. 55, 591–628.

Long, S.P., Ainsworth, E.A., Leakey, A.D.B., Nosberger, J., Ort, D.R., 2006. Food for thought: lower-than-expected crop yield stimulation with rising CO_2 concentrations. Science 312, 1918–1921.

Martin, R.V., Washington, R., Downing, T.E., 2000. Seasonal maize forecasting for South Africa and Zimbabwe derived from an agroclimatological model. J. Appl. Meteorol. 39 (9), 1473–1479.

Matsui, T., Horie, T., 1992. Effects of elevated CO_2 and high temperature on growth and yield of rice. 2. Sensitive period and pollen germination rate in high-temperature sterility of rice spikelets at flowering. Jpn. J. Crop Sci. 61, 148–149.

McMurtrie, R.E., Wang, Y.P., 1993. Mathematical models of the photosynthetic response of tree stands to rising CO_2 concentrations and temperatures. Plant Cell Environ. 16, 1–13.

McNaughton, K.G., Jarvis, P.G., 1991. Effects of spatial scale on stomatal control of transpiration. Agric. For. Meteorol. 54, 279–301.

Mearns, L.O., Easterling, W., Hays, C., Marx, D., 2001. Comparison of agricultural impacts of climate change calculated from the high and low resolution climate change scenarios. I. The uncertainty due to spatial scale. Clim. Change 51, 131–172.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculation by fast computer machines. J. Chem. Phys. 21, 1087–1092.

Mitchell, T.D., Jones, P.D., 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. Int. J. Climatol. 25, 693–712.

Monteith, J.L., 1995. Accommodation between transpiring vegetation and the convective boundary layer. J. Hydrol. 166, 251–263.

Morrison, J.I.L., 1987. Intercellular CO_2 concentration and stomatal response to CO_2. In: Zeiger, E., Farquhar, G.D., Cowan, I.R. (Eds.), Stomatal Function. Stanford University Press, Stanford, CA, pp. 229–251.

Morrison, J.I.L., Gifford, R.M., 1984a. Plant growth and water use with limited water supply in high CO_2 concentrations. 1. Leaf area, water use and transpiration. Aust. J. Plant Physiol. 11, 361–374.

Morrison, J.I.L., Gifford, R.M., 1984b. Plant growth and water use with limited water supply in high CO_2 concentrations. 2. Plant dry weight, partitioning and water use efficiency. Aust. J. Plant Physiol. 11, 375–384.

Neilson, R.P., 1995. A model for predicting continental-scale vegetation distribution and water balance. Ecol. Appl. 5, 362–386.

NRC (National Research Council), 1982. United State-Canadian Tables of Feed Composition. National Academy Press, Washington, DC.

Osborne, T.M., Lawrence, D.M., Challinor, A.J., Slingo, J.M., Wheeler, T.R., 2007. Development and assessment of a coupled crop-climate model. Global Change Biol. 13, 169–183.

Prentice, I.C., Sykes, M.T., Cramer, W., 1993. A simulation model for the transient effects of climate change on forest landscapes. Ecol. Model. 65, 51–70.

Prentice, I.C., Heimann, M., Sitch, S., 2000. The carbon balance of the terrestrial biosphere: ecosystem models and atmospheric observations. Ecol. Appl. 10, 1553–1573.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. Numerical Recipes in FORTRAN: The Art of Scientific Computing, 2nd ed. Cambridge University Press, New York, pp. 107–110.

Priestley, C.H.B., Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Mon. Weather Rev. 100, 81–92.

Qiu, J.J., Tang, H.J., Frolking, S., Boles, S., Li, C., Xiao, X., Liu, J., Zhuang, Y.H., Qin, X.G., 2003. Mapping single-, double-, and triple-crop agriculture in China at $0.5^\circ \times 0.5^\circ$ by combining county-scale census data with a remote sensing: derived land cover map. Geocarto Int. 18, 3–13.

Raupach, M.R., Rayner, P.J., Barrett, D.J., Defries, R.S., Heimann, M., Ojima, D.S., Quegan, S., Schmullius, C.C., 2005. Model-data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications. Global Change Biol. 11, 378–397.

Ritchie, J.T., Godwin, D.C., Otter-Nache, S., 1988. CERES-Wheat: A Simulation Model of Wheat Growth and Development. Texas A&M University Press, College Station, TX.

Rogers, H.H., Bingham, G.E., Cure, J.D., Smith, J.M., Surano, K.A., 1983. Responses of selected plant species to elevated carbon dioxide in the field. J. Environ. Qual. 12, 569–574.

Samarakoon, A.B., Gifford, R.M., 1996. Elevated CO_2 effects on water use and growth of maize in wet and drying soil. Aust. J. Plant Physiol. 23, 53–62.

Schlesinger, W.H., 1997. Biogeochemistry: An Analysis of Global Change. Academic Press, San Diego.

Sellers, P.J., Randall, D.A., Collatz, G.J., Berry, J.A., Field, C.B., Dazlich, D.A., Zhang, C., Collelo, G.D., Bounoua, L., 1996. A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: model formation. J. Clim. 9, 676–705.

Sionit, N., Rogers, H.H., Bingham, G.E., Strain, B.R., 1984. Photosynthesis and stomatal conductance with CO_2-enrichment of container- and field-grown soybeans. Agron. J. 76, 447–451.

Sitch, S., Smith, B., Prentice, I.C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J., Levis, S., Lucht, W., Sykes, M., Thonicke, K., Venevski, S., 2003. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic vegetation model. Global Change Biol. 9, 161–185.

Soil Conservation Service (SCS), 1991. Soil–plant–water relationships. In: National Engineering Handbook, Section 15 'Irrigation', Soil Conservation Service, U.S. Dept. of Agriculture, Washington, DC, 56 pp. (Chapter 1).

Spall, J.C., 2003. Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control. Wiley-Intersci. Ser. Discrete Math. Optim. Wiley-Interscience, Hoboken, NJ.

Tao, F., Yokozawa, M., Hayashi, Y., Lin, E., 2003. Changes in soil moisture in China over the last half-century and their effects on agricultural production. Agric. For. Meteorol. 118, 251–261.

Tao, F., Yokozawa, M., Zhang, Z., Xu, Y., Hayashi, Y., 2005. Remote sensing of crop production in China by production efficiency models: models comparisons, estimates and uncertainties. Ecol. Model. 183, 385–396.

Tao, F., Yokozawa, M., Xu, Y., Hayashi, Y., Zhang, Z., 2006. Climate changes and trends in phenology and yields of field crops in China 1981–2000. Agric. For. Meteorol. 138, 82–92.

Tao, F., Yokozawa, M., Liu, J., Zhang, Z., 2008a. Climate-crop yield relationships at province scale in China and the impacts of recent climate trend. Climate Res. 38, 83–94.

Tao, F., Hayashi, Y., Zhang, Z., Sakamoto, T., Yokozawa, M., 2008b. Global warming, rice production and water use in China: developing a probabilistic assessment. Agric. For. Meteorol. 148, 94–110.

Tarantola, A., 1987. Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation. Elsevier, New York.

Tjoelker, M.G., Oleksyn, J., Reich, P.B., 2001. Modelling respiration of vegetation: evidence for a general temperature-dependent Q_{10}. Global Change Biol. 7, 223–230.

Tubiello, F.N., Ewert, F., 2002. Simulating the effects of elevated CO_2 on crops: approaches and applications for climate change. Eur. J. Agron. 18, 57–74.

Tubiello, F.N., Soussana, J.F., Howden, S.M., 2007a. Crop and pasture response to climate change. Proc. Natl. Acad. Sci. U.S.A. 104, 19686–19690.

Tubiello, F.N., Amthor, J.S., Boote, K.J., Donatelli, M., Easterling, W., Fischer, G., Gifford, R.M., Howden, M., Reilly, J., Rosenzweig, C., 2007b. Crop response to elevated CO_2 and world food supply: a comment on ''Food for Thought'' by Long et al., Science 312:1918–1921, 2006. Eur. J. Agron. 26, 215–223.

Von Mises, R., 1964. Mathematical Theory of Probability and Statistics. Elsevier, New York.

Wheeler, T.R., Craufurd, P.Q., Ellis, R.H., Porter, J.R., Vara Prasad, P.V., 2000. Temperature variability and the annual yield of crops. Agric. Ecosyst. Environ. 82, 159–167.

Woodward, F.I., 1987. Climate and Plant Distribution. Cambridge University Press, New York.

Xie, P., Yatagai, A., Chen, M., Hayasaka, T., Fukushima, Y., Liu, C., Yang, S., 2007. A gauge-based analysis of daily precipitation over East Asia. J. Hydrometeorol. 8, 607–627.

Xu, T., White, L., Hui, D., Luo, Y., 2006. Probabilistic inversion of a terrestrial ecosystem model: analysis of uncertainty in parameter estimation and model prediction. Global Biogeochem. Cycles 20, GB2007, doi:10.1029/2005GB002468.

Zobler, L., 1986. A world soil file for global climate modelling. NASA Tech. Memorandum 87802, 32.