

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/304677884>

Uncertainties in predicting rice yield by current crop models under contrasting climatic environments, 2014.

Article · January 2015

CITATIONS

11

READS

951

24 authors, including:



Tao Li

International Rice Research Institute

52 PUBLICATIONS 2,904 CITATIONS

[SEE PROFILE](#)



Toshihiro Hasegawa

National Agricultural and Food Research Organization

225 PUBLICATIONS 7,717 CITATIONS

[SEE PROFILE](#)



Xinyou Yin

Wageningen University & Research

240 PUBLICATIONS 9,695 CITATIONS

[SEE PROFILE](#)



Yan Zhu

Nanjing Agricultural University

467 PUBLICATIONS 15,451 CITATIONS

[SEE PROFILE](#)

Uncertainties in predicting rice yield by current crop models under a wide range of climatic conditions

TAO LI^{1‡}, TOSHIHIRO HASEGAWA^{2‡}, XINYOU YIN^{3‡}, YAN ZHU^{4‡}, KENNETH BOOTE^{5‡}, MYRIAM ADAM⁶, SIMONE BREGAGLIO⁷, SAMUEL BUIS⁸, ROBERTO CONFALONIERI⁷, TAMON FUMOTO², DONALD GAYDON⁹, MANUEL MARCAIDA III¹, HIROSHI NAKAGAWA¹⁰, PHILIPPE ORIOL⁶, ALEX C. RUANE¹¹, FRANÇOISE RUGET^{8,16}, BALWINDER-SINGH¹², UPENDRA SINGH¹³, LIANG TANG⁴, FULU TAO¹⁴, PAUL WILKENS¹³, HIROE YOSHIDA¹⁰, ZHAO ZHANG¹⁵ and BAS BOUMAN^{1†}

¹International Rice Research Institute, Los Baños, Philippines, ²National Institute for Agro-Environmental Sciences, Tsukuba, Japan, ³Centre for Crop Systems Analysis, Wageningen University, Wageningen, The Netherlands, ⁴National Engineering and Technology Center for Information Agriculture, Nanjing Agricultural University, Nanjing, China, ⁵University of Florida, Gainesville, FL, USA, ⁶CIRAD, UMR AGAP, Montpellier, France, ⁷Cassandra Lab, DiSAA, University of Milan, Milan, Italy, ⁸UMR1114 EMMAH, INRA, Avignon F-84914, France, ⁹CSIRO Agriculture Flagship, Brisbane, Australia, ¹⁰National Agriculture and Food Research Organization, Tsukuba, Japan, ¹¹NASA Goddard Institute for Space Studies, New York, NY, USA, ¹²NASC Complex, CIMMYT 110008, New Delhi, India, ¹³International Fertilizer Development Center, Muscle Shoals, AL, USA, ¹⁴Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China, ¹⁵State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China, ¹⁶UMR1114 EMMAH, UAPV, Avignon F-84914, France

Abstract

Predicting rice (*Oryza sativa*) productivity under future climates is important for global food security. Ecophysiological crop models in combination with climate model outputs are commonly used in yield prediction, but uncertainties associated with crop models remain largely unquantified. We evaluated 13 rice models against multi-year experimental yield data at four sites with diverse climatic conditions in Asia and examined whether different modeling approaches on major physiological processes attribute to the uncertainties of prediction to field measured yields and to the uncertainties of sensitivity to changes in temperature and CO₂ concentration [CO₂]. We also examined whether a use of an ensemble of crop models can reduce the uncertainties. Individual models did not consistently reproduce both experimental and regional yields well, and uncertainty was larger at the warmest and coolest sites. The variation in yield projections was larger among crop models than variation resulting from 16 global climate model-based scenarios. However, the mean of predictions of all crop models reproduced experimental data, with an uncertainty of less than 10% of measured yields. Using an ensemble of eight models calibrated only for phenology or five models calibrated in detail resulted in the uncertainty equivalent to that of the measured yield in well-controlled agronomic field experiments. Sensitivity analysis indicates the necessity to improve the accuracy in predicting both biomass and harvest index in response to increasing [CO₂] and temperature.

Keywords: AgMIP, climate change, crop-model ensembles, *Oryza sativa*, yield prediction uncertainty

Received 14 June 2014; revised version received 11 September 2014 and accepted 12 September 2014

Introduction

Rice (*Oryza sativa*) is the staple food for nearly half of the world's population and its productivity under future climates is crucial for global food security (Godfray *et al.*, 2010). Rice is grown under a wide range of

environments, from the humid tropics to the temperate zones, and the impact of climate change is likely to differ regionally (Wassmann *et al.*, 2009).

The impact of climate change on crop production has been assessed using ecophysiological process-based crop models combined with climate models (Tao *et al.*, 2008). However, different crop models are used in different studies, and uncertainties associated with such crop models are rarely assessed. This is in contrast to the systematic use of a multiple number (or 'ensemble') of models commonly practiced in the estimation of the direction, magnitude, and uncertainties of future climate change (Meehl *et al.*, 2007a,b). A comprehensive

Correspondence: Toshihiro Hasegawa, tel. +81 29 838 8204, fax +81 29 838 8211, e-mail: thase@affrc.go.jp; X. Yin, tel. +31 317 482 348, fax +31 317 485 572, e-mail: Xinyou.Yin@wur.nl

[†]The leader of AgMIP Rice Team.

[‡]Members of leading group of AgMIP Rice Team. All other authors had equivalent contributions and are in alphabetical order by surnames.

review on simulated future crop yields was presented by Easterling *et al.* (2007). In general, increasing temperatures are projected to have a negative effect on yield, but elevated carbon dioxide concentration [CO₂] will attenuate the reduction by promoting photosynthesis and biomass production, and the magnitude of the overall effects is uncertain and depends on the regions (varying from low to mid to high latitudes). Unfortunately, however, the sources of uncertainty were not identified because individual simulations were taken from independent studies at different sites with various growth and management conditions.

The Agricultural Model Intercomparison and Improvement Project (AgMIP) was launched in 2010 to assess the impacts of climate change on agricultural production and food security using process-based crop models along with climate and economic simulations (Rosenzweig *et al.*, 2013). The Agricultural Model Intercomparison and Improvement Project also emphasizes efforts to quantify the sources of uncertainty in yield prediction for major crops, and the identification of options to reduce that uncertainty. Within AgMIP activities, a protocol for crop model comparison has been proposed and recently applied to wheat models (Asseng *et al.*, 2013) and maize models (Bassu *et al.*, 2014), with the models considerably differing in their prediction of wheat or maize yield response to climate factors. Unlike wheat and maize, most rice is produced in submerged paddy conditions, and associated uncertainties may be different from temperate upland crops. We, the AgMIP Rice team, used 13 models to estimate uncertainties in yield and biomass predictions and to test the hypothesis that an ensemble of models reduces uncertainty. Simulations were made for irrigated paddy conditions as this system produces 75% of the total rice production (Maclean *et al.*, 2002). All models were evaluated against field experiments and regional yield records at four sentinel sites in different ecological zones in Asia.

Materials and methods

Datasets

Four sentinel sites, Los Baños (Philippines), Ludhiana (India), Nanjing (China), and Shizukuishi (Japan) were chosen to represent contrasting rice-growing areas in Asia (Table 1). Six to eight field experiments at each site were available for model evaluation. From each site, only the treatments with full irrigation and fertilizer supply aimed at potential crop growth were used for calibration (see details in Simulation analysis section). In addition, regional statistical yield data for 1980–2009 were collected from local government publications, which can be retrieved from the

following sources: for Los Baños, the GIS research group, Social Sciences Division, International Rice Research Institute, Philippines; for Ludhiana, Statistical Abstract of India, Ministry of Agriculture, Government of India and Statistical Abstract of Punjab, Department of Agriculture, Government of Punjab; for Nanjing, Statistical Yearbook of Nanjing at <http://www.njtj.gov.cn/2004/>; and for Shizukuishi, Annual rice production statistics by municipality, Japan, and Ministry of Agriculture, Forestry and Fisheries, (1980–2010), data available at http://www.maff.go.jp/j/tokei/kouhyou/sakumotu/sakkyou_kome/index.html#c. Daily weather data during the experiments, and for baseline years 1980–2010, were obtained from weather stations on-site.

Rice models and their classification

Thirteen models differing in structure, forcing variables and input parameters were used (Table S1 and S2, hereafter, 'S' before table/figure numbers is used to indicate supplementary material). These models cover almost all the major activities of modeling field-level performance of rice growth and yield in response to variable climatic conditions. They commonly use daily climatic factors such as solar radiation and temperature as their input variables. Twelve of the 13 models consider crop response to [CO₂] (Table S2). All models provide at least three fundamental output variables; growth duration, biomass and grain yield. The 13 rice models can be classified according to how they model the following three major subprocesses that are highly relevant in determining yield responses to climate changes.

Phenology or growth duration. All models use temperature as a major driver for the progress of developmental stages (Table S2), but functional forms for the developmental rate are different. Models were grouped into the following two categories (Table S2): those that use a linear or nonlinear temperature response function with an upper or asymptotic limit for the developmental rate, and those with a function with an optimum temperature above which the developmental rate decreases with the temperature increase.

Net primary or biomass production. Net primary production is driven typically by solar radiation and [CO₂], and in some models is also moderated by temperature (Table S2). But models differ in complexity and/or empiricity. For net primary production, we grouped the 13 models into three categories: models based on the canopy radiation use efficiency (RUE) (Monteith & Moss, 1977), those based on the light-response curve of single-leaf photosynthesis and its integration to the canopy (LRC) (De Wit, 1978), and those based on the Farquhar-von Caemmerer-Berry biochemical model of leaf photosynthesis and its integration to the canopy (FvCB) (Farquhar *et al.*, 1980).

Spikelet fertility in response to temperature. Yield sensitivity to temperature can vary substantially, depending on whether models account for spikelet sterility, which is one of the major

Table 1 Details of the sentinel sites and the field experiments

Dataset location	Los Baños	Ludhiana	Nanjing	Shizukuishi
Institute	IRRI*	PAU†, IRRI	NAU‡	NIAES, NARO/ TART§
Site conditions				
Country	Philippines	India	China	Japan
Geo-coordinates	121°9'E, 14°6'N	75°48'E, 30°54'N	118°59'E, 32°56'N	140°57'E, 39°41'N
Climate type	Tropical, humid	Subtropical, semiarid	Subtropical, semihumid	Cool temperate, humid
Annual precipitation (mm)	2007	752	1090	1547
Annual mean of daily radiation (MJ m ⁻² day ⁻¹)	16.5	18.2	11.8	12.0
Annual mean of daily maximum temperature (°C)	30.7	29.5	20.6	13.9
Annual mean of daily minimum temperature (°C)	23.0	16.3	12.1	5.0
Years of experiments	1985&1986	2008 & 2009	2007	1998, 1999 & 2000
Treatments	2 seasons with 3-4 nitrogen (N) levels (7 treatments)	2 seasons with 4 irrigation regimes (8 treatments)	1 season with 2 cultivars and 3 N levels (6 treatments)	3 seasons with 2-3 N levels (8 treatments)
Rice field management				
Crop establishment	Transplanting	Transplanting	Transplanting	Transplanting
Water management	Flooding	Flooding or AWD ⁵	Flooding	Flooding
Nitrogen fertilizer management	30, 60, 90 & 120 kg ha ⁻¹ in 1985; 35, 70 & 140 kg ha ⁻¹ in 1986	120 kg ha ⁻¹	110, 220 & 330 kg ha ⁻¹	80 & 120 kg ha ⁻¹ in 1998; 40, 90 & 150 kg ha ⁻¹ in 1999 and 2000.
Rice genotypes used				
Cultivar name	IR58	PAU201	Wuxiangjing 14 (V1) and Liangyoupeiiju (V2)	Akitakomachi
Ecotype	Indica	Indica	V1, Japonica; V2 Indica hybrid	Japonica
Growth duration	120–125 days	130–135 days	ca 150 days (V1); ca 145 days (V2)	ca 150 days
Available measurements				
Biomass of organs and total above-ground biomass	Yes	Yes	Yes	Yes
Nitrogen of organs	Yes	No	Yes	Yes
Leaf area index	No	Yes	Yes	Yes
Grain yield	Yes	Yes	Yes	Yes
Phenology	Yes	Yes	Yes	Yes
Treatment used for model calibration**††‡‡	1985-N30, 60, 90 & 120	2008-MT, MT1, MT2 & MT3	2007-V1-N330 & V2-N330	1999-N40, 90 & 150
Treatment for sensitivity analysis**††‡‡	1985-N120	2008-MT, N120	2007-V1-N220	1999-N150

Table 1 (continued)

Dataset location	Los Baños	Ludhiana	Nanjing	Shizukuishi
Contact person for detail of dataset	Roland Buresh	Sudhir-Yadav	Yan Zhu	Toshihiro Hasegawa

*IRRI = International Rice Research Institute, Los Baños, Philippines.

†PAU = Punjab Agricultural University, Ludhiana, India (Sudhir-Yadav *et al.*, 2011a, b).

‡NAU = Nanjing Agricultural University, Nanjing, China (Unpublished dataset, please contact data owner if detail is needed).

§NIAES = National Institute for Agro-Environmental Sciences, Tsukuba, Japan; NARO/TART = National Agricultural Research Organization/Tohoku Agricultural Research Center (Kim *et al.*, 2003).

**The initial four-digit number stands for the year. MT stands for the soil moisture treatments; Continuous flooding (MT), alternate wet and dry (AWD) field water management with irrigation applied to have 50 mm depth of surface water when the soil water tension at 20 cm increased to 20 kPa (MT1), at 40 kPa (MT2) and at 70 kPa (MT3). All treatments were kept continuously flooded at 50 mm standing water for the first 15 DAT prior to commencement of the irrigation treatments.

††Values followed by N is the rate of N application in kg ha⁻¹.

‡‡V1 and V2 are the varieties, and their names are given in the genotypes descriptions in this table.

sources of yield variation under extreme conditions, as affected by heat or cold. We therefore grouped the 13 models, according to whether or not the models account for temperature-induced changes in spikelet sterility (Table S2). It is expected that spikelet sterility greatly decreases crop sink size and harvest index.

Simulation analysis

Four sets of simulations were created and implemented by the 13 crop models:

In the *'blind' simulation*, for each site, each model used its own default set of parameters for rice with minimum calibration only for crop phenology to estimate rice growth and yield under both ambient [CO₂] (NOAA ESRL data) with on-site weather and [CO₂] of 360 μmol mol⁻¹ with the historic 30-year baseline weather conditions.

In the *calibration and validation simulation*, additional experimental data on crop yield and time courses of leaf area index and biomass of crop organs were provided from a subset of the experiments (specified in Table 1 and Figure S1), to allow full calibration of the parameters of each model. Individual modelers were requested to choose among four methodologies for this calibration (Table S3), and the differences in the calibrated parameter values from the *'blind' simulation* stage are given in Table S3. The calibrated models were then evaluated against the remaining independent dataset of the experiments (Table 1).

In the *sensitivity analysis*, the models with calibrated parameters were run for each site with the crop management settings of an experimental treatment presented in Table 1 with the historic 30-year weather data and [CO₂] of 360 μmol mol⁻¹ as baseline. The sensitivity to [CO₂] was simulated at two temperature conditions (0 and +3 °C increase) by varying [CO₂] from baseline to 720 μmol mol⁻¹ with an incremental step of 90 μmol mol⁻¹. The sensitivity to temperature was analysed in the baseline and doubled [CO₂] condition, by changing the air temperature from -3 to +9 °C in 3 °C intervals [for [CO₂] = 360 μmol mol⁻¹] or by increasing air temperature by 3 and 6 °C (for [CO₂] = 720 μmol mol⁻¹).

In the *climate change simulation*, the calibrated models were run at each site using the crop management settings of the treatments in the field experiment and the 30-year midcentury (2040 to 2069) temperature projections with 16 global climate model (GCM)-based sub-scenarios under the A2 emission scenario reported in the SRES (Special Report on Emission Scenario) (Easterling *et al.*, 2007). The sub-scenarios were based on the outputs of 16 GCMs for the midcentury from the CMIP3 (Third Coupled Model Intercomparison Project) Multi-model Dataset (Meehl *et al.*, 2007a). Specifically, we firstly calculated monthly temperature increments between the 1980–2009 and 2040–2069 from each scenario. These increments were added to the measured daily minimum and maximum temperature data for 1980–2009 according to month, while other climatic factors (such as solar radiation and relative humidity) were not modified. Thus, the future weather data used in this study had the same magnitude of interannual and intraseasonal variation as that of the baseline weather data. Details of these procedures are given in the AgMIP Climate Team protocols (Rosenzweig *et al.*, 2013). [CO₂] was common across 16 GCM-based climate scenarios and held constant at 556 μmol mol⁻¹, which is the predicted [CO₂] for the midtime point of the period between 2040 and 2069 under the A2 emission scenario.

All simulated and measured yield refers to rough rice yield (i.e., the unmilled grain plus the hull) at 14% moisture, whereas biomass refers to the total above-ground biomass of the whole crop at the time of harvest at 0% moisture content.

Data analysis

For all models, simulated yield and biomass of the *'blind' simulation* as well as the *calibration and validation simulation* were compared with field observations, and goodness-of-fit was assessed using common statistic metrics (Table S4). Interannual variability in simulated yield using the calibrated models and the historic weather data under [CO₂] of 360 μmol mol⁻¹ was compared with the variability in regional yield data.

Regional yields are affected by both year-to-year climatic conditions and long-term changes in climate and technologies.

To evaluate year-to-year climatic effect on simulated yields based on the fixed cultivation management and constant [CO₂] with observed yields, we firstly extracted the trend component in the observed yields, assuming that this reflects the effects of technological improvement and [CO₂] elevation over the years. Then we corrected the observed yields with the trend component to estimate de-trended yields.

The specific de-trending procedures are as follows; (i) developing polynomial regression function for each site, using data of original regional yield, Y_i^o of 30 years from 1980 (excluding the data recorded in abnormal weather years from the analysis); (ii) estimating annual yield (Y_i') using the regression equations (Figure S2); (iii) calculating the mean value (Y_x) of Y_i' over the 30 years; (iv) deriving the difference as $D_i = Y_i' - Y_x$; and (v) calculating de-trended yield as $Y_i = Y_i^o + D_i$.

Interannual variability was quantified by use of yield deviation (or anomaly), which was calculated for a given year (i) during the 1980–2009 period as $\frac{Y_i - \bar{Y}}{\bar{Y}} \times 100$ where Y_i is the simulated yield for year i and \bar{Y} is the average of all Y_i over 30 historical years. For simulations under future climate, Y_i is the average simulated yield under 16 sub-scenarios for year i and \bar{Y} is the average of all Y_i over 30 future years. For regional yield anomaly, Y_i is the de-trended regional yield in year i and \bar{Y} is the average of all Y_i over a 30-year period of statistical data.

To examine how an ensemble of multiple models affects the uncertainty of predicted yield, we first computed mean of predicted yields with different numbers of ensemble models, ranging from 1 to 13 [Eqn (1)];

$$\bar{P}_{n,j} = \frac{\sum_{i=1}^n P_{ij}}{n} \quad (n = 1, 2, 3, \dots, 13) \quad (1)$$

where P_{ij} is the predicted yield by model i for the experiment j and $\bar{P}_{n,j}$ is the average of predictions by n models for the

experiment j . In this study, we defined uncertainty (U_n , %) as the ratio between the summation of the absolute deviation of $\bar{P}_{n,j}$ from measured yield of experiment j , O_j , and the summation of the experimental yields [Eqn (2)];

$$U_n = \frac{\sum_{j=1}^k |\bar{P}_{n,j} - O_j|}{\sum_{j=1}^k O_j} \times 100 \quad (2)$$

where k is the number of experiments used for the evaluation. For each number of model ensemble (n), we calculated U_n for all possible combinations of models (i.e. $C_{13}^n = \frac{13!}{n!(13-n)!}$ times) to determine the range of uncertainty at a given number of models.

Results

Model performance in representing field measurements

In the low-calibrated 'blind' simulation state, the ensemble mean and median of predicted yield and biomass from all models were mostly within the 95% confidence interval of measurements (Fig. 1, label M). While the difference between the means of predictions of all models and the means of the measurements reached up to 17%, the yield predictions of individual models differed as much as twofold, beyond the 95% confidence limits of measured yields (Fig. 1a–e). The ensemble mean of predicted yield in blind and calibration simulations in each treatment was also similar to the measured yield and biomass, which differed substantially among the treatments, sites and years (Figure S1). Predictions and metrics for goodness of fit also differed substantially among the models (Table S4). After thorough

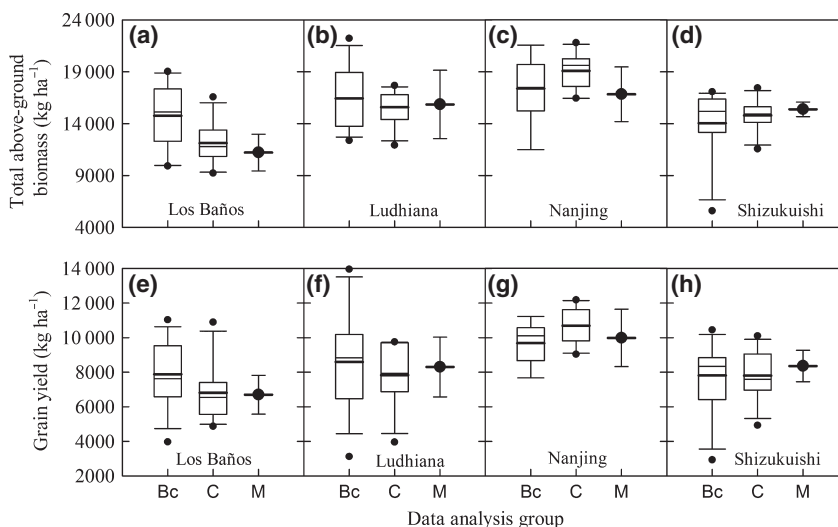


Fig. 1 Measured and simulated above-ground biomass (a–d) and grain yield (e–h) of rice for 13 models at the four sentinel sites. Bc, 'blind simulation' (calibrated for phenology only); C, simulation after detailed model calibration; M, measured values with 95% confidence intervals based on 3 to 4 replicated measurements of treatments in each site. The boxes present the 25-to-75 percentile, whiskers (i.e., the horizontal dashes linked to boxes by vertical lines) present the 10-to-90 percentile, thick and thin dashes in boxes are mean and median of all model prediction, and black dots are the simulated results outside the 10-to-90 percentile.

calibration, model performance improved significantly (Fig. 1, label C), and the difference between the means of predictions and measurements was less than 10% in all cases.

Simulation of historic and future rice yield

The interannual variability in yield simulated by the models for baseline of 30 years (1980–2009) showed good agreement with the de-trended regional yields at all four sites if averaged over the 13 models (Fig. 2a–d). The yearly deviation of the simulated yields significantly correlated with that of the de-trended

yields ($r = 0.422$, $P < 0.001$), with the root mean square error 8.9% for the observed deviations ranging from -79.3% to 16.9% (Fig. 2). Variability in the observed yields was largest at Shizukuishi, where the growing-season air temperature was lowest ($18.5\text{ }^{\circ}\text{C}$). Year-to-year variation at Shizukuishi was predicted to be reduced under future climate with higher temperatures ($20.2\text{ }^{\circ}\text{C}$, Fig. 2h).

We also assessed rice production under the midcentury (2040–2069) climates, based on GCM-based scenarios generated from the outputs of 16 climate models under a relatively high emission scenario (A2; SRES, 2000). The relative change in grain yield under midcen-

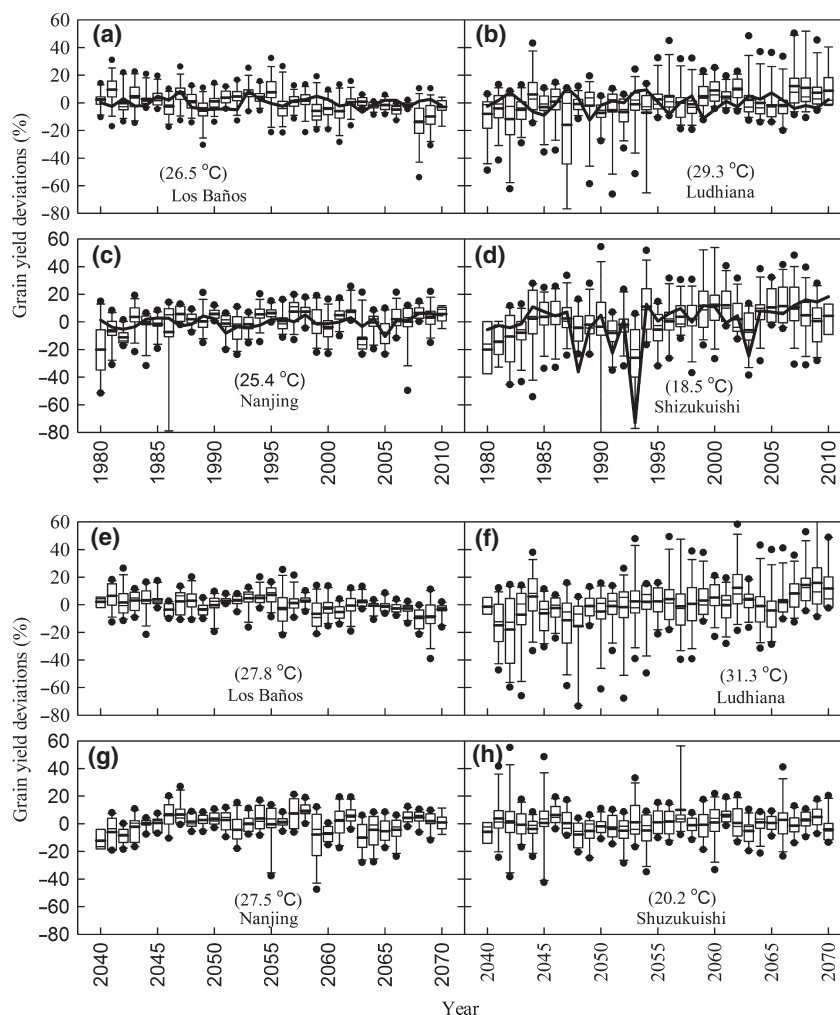


Fig. 2 Predicted and regional statistical inter-annual variability in rice yield. Thick solid lines in panels a–d are the annual variability (deviation from 30-year mean) derived from de-trended regional yield data. Boxes and whiskers in all panels present the variation in predicted yield variability of 13 rice models for individual years under historical climatic conditions (1980 to 2010) with atmospheric $[\text{CO}_2]$ of $360\ \mu\text{mol mol}^{-1}$ (a–d), and under projected future climatic conditions (2040 to 2069, 16 GCM-based scenarios) with atmospheric $[\text{CO}_2]$ of $556\ \mu\text{mol mol}^{-1}$ (e–h), at the four sentinel sites. Boxes and whiskers present 25 to 75 and 10 to 90 percentiles of variability, respectively, among the models. The thick and thin dashes in boxes are the mean and median of predicted relative yield changes, respectively. The black dots are the variability beyond the 10 or 90 percentiles. The values in parentheses are the average air temperature in crop growth period over 30 years.

ture climates from the baseline varied from -46% to 80% among sites and models (Fig. 3). The different crop models predicted modest changes in yield (-8% to 28%) at Los Baños and Nanjing (Fig. 3a, c), but substantial differences among the crop models were observed at the other two sites. The crop models did not agree even in the direction of yield change at Ludhiana, ranging from -51% to $+78\%$ (Fig. 3b). On the other hand, mean yield changes over the crop models were similar across different climate scenarios, except for the warmest site, Ludhiana.

Modeled response to rising CO_2 and temperature

With the increases in temperature and $[\text{CO}_2]$ from current baseline values at each site, predicted changes in yield varied by model, site, and magnitude of change in temperature and $[\text{CO}_2]$ (Fig. 4). In general, however, predicted yield decreased with increasing temperature and increased with increasing $[\text{CO}_2]$ at all sites. Also, the variation among model estimates increased as temperature and $[\text{CO}_2]$ moved further from observed conditions. The responses to temperature increases were site-specific, but were little influenced by $[\text{CO}_2]$ [Fig. 4a–h, temperature increases from -3 to $+9$ °C for a comparison between two $[\text{CO}_2]$ levels]. Yields decreased by 1% to 10% per degree temperature increase under $[\text{CO}_2]$ of $720 \mu\text{mol mol}^{-1}$, and by 2% to 11% under $[\text{CO}_2]$ of $360 \mu\text{mol mol}^{-1}$ (Table 2). The decrease in yield with increasing air temperature was

highest at Ludhiana as this site experienced the highest temperatures in the current conditions (Fig. 4).

Yield responses to $[\text{CO}_2]$ were higher in the $360\text{--}450 \mu\text{mol mol}^{-1}$ range (7% to 11% per $100 \mu\text{mol mol}^{-1}$) than in the $630\text{--}720 \mu\text{mol mol}^{-1}$ range (3% to 5% per $100 \mu\text{mol mol}^{-1}$) at all sites, indicating a diminishing return from increasing $[\text{CO}_2]$. The responses to increases in $[\text{CO}_2]$ were hardly affected by temperature at Los Baños, Nanjing, and Shizukuishi. At Ludhiana, the yield increases were 2% to 4% higher per $100 \mu\text{mol mol}^{-1}$ $[\text{CO}_2]$ increase at current temperature than with a 3 °C temperature increase (Fig. 4j, n).

The substantial variation in Figs 3 and 4 may be explained by differences in the model approaches used to reproduce temperature and $[\text{CO}_2]$ effects on crop phenology, net primary production and spikelet fertility.

Response of growth duration to the changes in temperature was different among models and sites (Figure S3). All models predicted a shorter growth duration with the increases in temperature at Shizukuishi, but at the other three sites, models disagreed; some predicted a shorter duration whereas others predicted a prolonged duration as the temperature rose by more than 3 °C. As expected, models with an upper or asymptotic limit to the developmental rate at high temperatures (red dots in Figure S3) consistently showed a shorter duration with temperature increases. Substantial variation occurred in the upper temperature range among the models having an optimum temperature

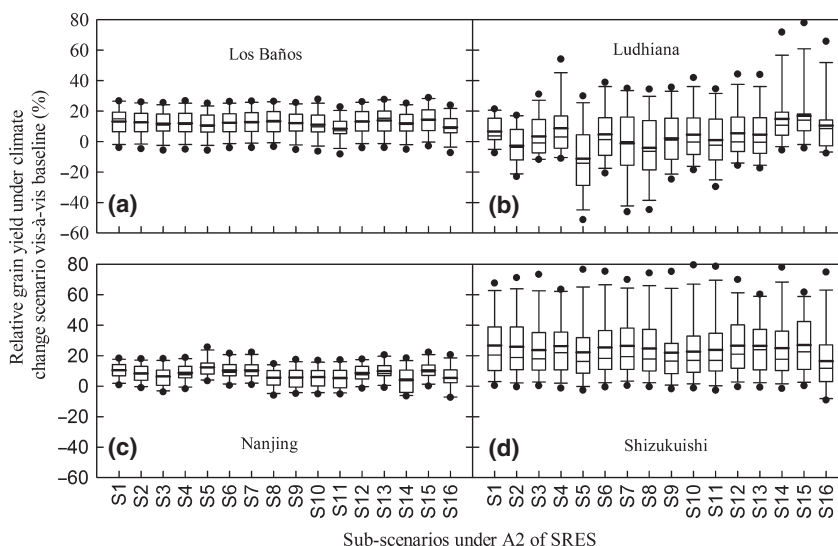


Fig. 3 Relative change in simulated grain yields by models under 16 GCM-based climate scenarios under the A2 emission scenario (S1 to S16). The projected changes of annual mean temperature at each site, which were averaged across all climate models ranged from 1.3 to 2.1 °C (Figure S4). The average grain yield of 30-year (2040–2069) simulations was used to develop the change in grain yield for each scenario relative to the average simulated yield using baseline years in individual models. The boxes and whiskers present the variation in relative change among models. The thick and thin dashes in boxes are mean and median of the relative change in all model predictions.

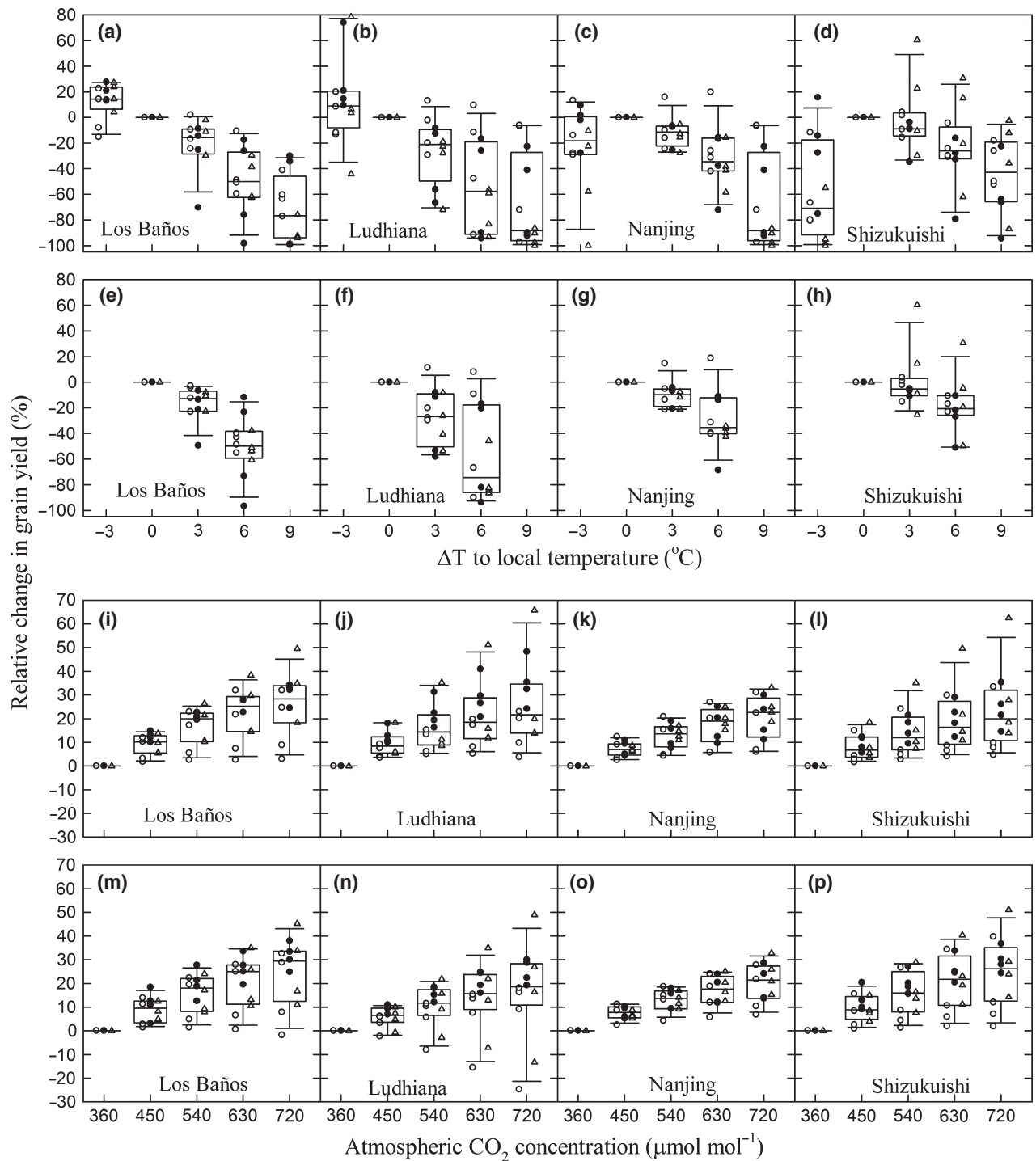


Fig. 4 Relative change in predicted rice yield with variation in temperature and in atmospheric $[CO_2]$ at the four sentinel sites. a to d, response to temperature increase from local current mean at $[CO_2]$ of $360 \mu mol mol^{-1}$; e to h, at $[CO_2]$ of $720 \mu mol mol^{-1}$; i to l, response to $[CO_2]$ increase from $360 \mu mol mol^{-1}$ to $720 \mu mol mol^{-1}$, at current local air temperatures; m to p, at temperature increased by $3 \text{ }^\circ C$. Boxes and whiskers present 25-to-75 and 10-to-90 percentiles of models, respectively. The dashes in boxes are the median of predicted relative yield changes of all models. The black dots are for models with the FvCB approach for net primary productivity, open-triangles are for models with the RUE approach, and open-circles are for models using the LRC approach (for model classification, see the main text). (Note: The model SAMARA does not have CO_2 response programmed in, and was therefore excluded in the analysis of CO_2 -by-temperature interaction.)

Table 2 The average change rates of grain yields with temperature increases and increases in CO₂ derived from all model predictions

Item	Los Baños	Ludhiana	Nanjing	Shizukuishi	Average
Decrease (%) per degree increase under 360 $\mu\text{mol mol}^{-1}$					
With 0 to 3 °C temperature increase	6.65	9.10	3.88	1.61	5.31
With 3 to 6 °C temperature increase	9.60	9.32	6.63	7.69	8.31
Decrease (%) per degree increase under 720 $\mu\text{mol mol}^{-1}$					
With 0 to 3 °C temperature increase	5.50	9.11	3.33	0.60	4.64
With 3 to 6 °C temperature increase	11.06	9.54	6.56	6.84	8.50
Increase (%) per $\mu\text{mol mol}^{-1}$ of increase in CO ₂ with no temperature change					
With 360 to 450 $\mu\text{mol mol}^{-1}$ CO ₂	0.10	0.10	0.08	0.09	0.09
With 450 to 540 $\mu\text{mol mol}^{-1}$ CO ₂	0.08	0.08	0.06	0.07	0.07
With 540 to 630 $\mu\text{mol mol}^{-1}$ CO ₂	0.06	0.06	0.05	0.05	0.06
With 630 to 720 $\mu\text{mol mol}^{-1}$ CO ₂	0.04	0.05	0.04	0.04	0.04
Increase (%) per $\mu\text{mol mol}^{-1}$ of increase in CO ₂ with 3 °C temperature increase					
With 360 to 450 $\mu\text{mol mol}^{-1}$ CO ₂	0.10	0.07	0.08	0.11	0.09
With 450 to 540 $\mu\text{mol mol}^{-1}$ CO ₂	0.08	0.05	0.06	0.07	0.06
With 540 to 630 $\mu\text{mol mol}^{-1}$ CO ₂	0.06	0.04	0.04	0.06	0.05
With 630 to 720 $\mu\text{mol mol}^{-1}$ CO ₂	0.05	0.03	0.04	0.04	0.04

function, in which development slows down at temperatures above the optimum (black dots). The phenology response to temperature under elevated [CO₂] was essentially similar to that under current [CO₂] (Figure S3a–d vs. e–h) because all the models assume no impact of [CO₂] on crop phenology.

All models predicted that crop biomass (Figure S4) and grain yield (Fig. 4) decreased with increasing temperature and increased with rising [CO₂], and variation among models increased with the increase in temperature and [CO₂]. Despite different model structures for net primary production, the models did not show a distinct 'structure-related' pattern for the biomass response to temperature (Figure S4a–h), or for the response to [CO₂] (Figure S4i–p). However, FvCB-type (Table S2) models (black dots in Fig. 4) tended to show smaller variation in the response to [CO₂] among models and across sites, whereas RUE-type models (open-triangles) showed larger variation among models and across sites (Fig. 4i–p). Most RUE-type models showed a limited response to [CO₂] at Shizukuishi (Figure S4i, p), but a stronger response at the other sites. The variation among models within the LRC-type (open-circles) was moderate.

Harvest index (HI) was highly sensitive to changes in temperature (Figure S5a–h). Most models predicted a decreasing HI with rising temperature, with a few exceptions at the cooler Shizukuishi site, in which some models predicted an increase in HI (Figure S5d, h). Apparently, models that account for heat and/or cold-induced sterility (red dots) showed a higher sensitivity of HI to temperature. Variation among models was particularly large at higher temperature ranges at Los Baños and Ludhiana, and at low temperature (–3 °C)

at Shizukuishi. [CO₂] had a limited effect on HI compared to temperature under current climate (Fig. S5i–p), but models tended to predict a decrease in HI with increasing [CO₂]. The decrease in HI with rising [CO₂] was particularly pronounced in a few models under a +3 °C condition at Los Baños and Ludhiana (Figure S5m, n).

Ensemble model prediction of yield

Ensemble modeling significantly decreased the uncertainty of projecting measured yield at the four sentinel sites (Fig. 5a, b). The means of predictions of all models reduced uncertainty to about 5% of measured yields (Fig. 5b), whereas any single model did not provide predictions with uncertainties of less than 10% across all locations. The uncertainty of prediction decreased rapidly with the increasing number of models used for ensemble predictions. Variation in measured yield, caused by variation among replicates, sampling errors, and measurement errors in agronomic field experiments, is typically around 15% (Belder *et al.*, 2007). To obtain a similar uncertainty in model prediction, at least eight low-calibrated models or five well-calibrated models were needed.

Discussion

Confidence in predicting rice yield

Model performance showed the limited capability of individual models to consistently reproduce yield over different experimental conditions (Fig. 1). Calibration with detailed crop growth data from a

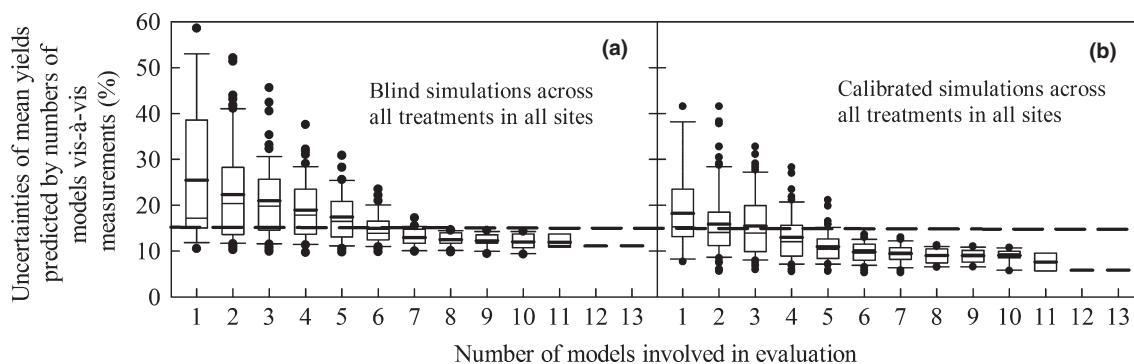


Fig. 5 Uncertainty of mean rice yield predicted by a different number of models. The uncertainties were expressed as the deviation (%) of the mean simulated yield across the models from the experimentally measured yield (a, b). Boxes and whiskers present 25-to-75 and 10-to-90 percentile ranges of uncertainty, respectively. The thick and thin dashes in boxes are the mean and median of prediction uncertainty, respectively. The black dots are relative changes beyond the 10-to-90 percentile. The broken lines in panels a and b indicate the 15% uncertainty presumed representative of typical field measurements (Belder *et al.*, 2007).

subset of the experiments reduced the range of 10 to 90 percentile of yields predicted among models by 48% to 57% for total biomass (Fig. 1a–d) and 7% to 46% for grain yield (Fig. 1e–h). A similar reduction in model-to-model variation due to calibration was reported for the wheat and maize uncertainty evaluations (Asseng *et al.*, 2013; Bassu *et al.*, 2014), but the magnitude of the reduction was less in our study than in the case of wheat or maize. The reasons for this are not clear, but even with calibrated models, predicted yields had a larger variation among models compared to 95% confidence interval of the experimental data, except at Nanjing. On the other hand, the mean of 13 models, even at the blind simulation stage (except for Los Baños where the uncertainty of predicted grain yield was 17%), represented the experimental yield with less than 10% uncertainty for all climate conditions (Fig. 1).

The mean of the 13 calibrated models also showed good agreement with observed trends in the interannual yield variability in the 30-year historical yield data, suggesting that current models can collectively capture yield fluctuations over 30 years (Fig. 2a–d). The models used in this study take into account only the environmental effects of temperature, radiation, humidity, and [CO₂], and were run under scenarios of full irrigation and adequate access to nutrients. In practice, farmers may experience effects of pests, diseases, weed infestations, and possible suboptimal management of water and nutrients, which all affect crop growth and yield formation. Nevertheless, the good agreement between modeled yield variability and de-trended regional yield variability obtained by farmers suggests that the models can capture a substantial amount of yield variability caused by environments and inter-annual variation in temperature and radiation.

Model uncertainty increases at both low and high temperatures

Differences among the models in the predicted yields and 30-year historical yield variability were noted at the coolest (Shizukuishi) and warmest (Ludhiana) sites (Fig. 1e–h, Fig. 2a–d), suggesting that uncertainty increases at both low and high temperatures. The increased variability in predicted yield among the models with increases in temperature in the sensitivity analysis supports this observation for the high-temperature sites (Fig. 4a–h). This variability is large because the models differ substantially in the way, and in the extent to which, extreme temperature is modeled to affect crop phenology and yield formation, particularly grain setting and/or harvest index (Zhang & Tao, 2013). At Shizukuishi, the variation among the crop models was smaller for biomass predictions than for grain yield predictions (Fig. 1d, h). Cool spells in midsummer occur frequently in northern Japan; low temperatures at the microspore stage result in cold-induced sterility, thereby substantially reducing yields. Overall, reproductive growth of rice is highly sensitive to both high- and low-temperature extremes (Satake, 1976; Satake & Yoshida, 1978; Jagadish *et al.*, 2007), and uncertainty in the predictive power of crop models in capturing the effects of temperature extremes is still large (Hasegawa *et al.*, 2011).

Variation among the models in future yield projections is site-dependent

Climate data for the midcentury (2040–2069) were generated from outputs of 16 climate models under a relatively high emission scenario (A2; SRES, 2000). Nevertheless, variation among crop models in the prediction of inter-annual yield variation increased from past 30-year his-

toric values at Ludhiana, remained unchanged at Los Baños and Nanjing, and decreased at Shizukuishi (Fig. 2e–h). The decrease in inter-model variation at Shizukuishi resulted from the reduced occurrences of cool spells under future climates, whereas, at Ludhiana, extreme high temperatures in certain years under future climates induced severe yield reductions in some models, but not in all (Figs 3 and 4).

Site-dependent variation was also underlined in the yield change relative to the baseline climatic conditions (Fig. 3). The crop models predicted a modest decrease and similar yield response (–8% to 28%) at Los Baños and Nanjing (Fig. 3a, c), but substantial differences among the crop models were observed at the other two sites. In particular, predicted yield change at Ludhiana ranged from –51% to +78% among the models (Fig. 3b). It is worth noting that the variation in temperature increases among climate subscenarios was similar across different sites. This suggests that the large variation in yield prediction due to climate subscenarios at Ludhiana was not because of a larger variation in temperature increases but because of the higher sensitivity of crop models to elevated temperature. Globally, climate projections generally agree that the projected temperature increase is smaller at low latitudes than high latitudes (Meehl *et al.*, 2007b). Given relatively small increases in temperature, however, uncertainty in yield predictions could be amplified by higher yield sensitivity to temperature at low latitude where the baseline temperature is already high as was seen by Rosenzweig *et al.* (2014). Improvements in both crop models and climate models are strongly needed for these regions.

Sources of uncertainties

The large variation among models in simulated yield responses to increases in temperatures and [CO₂] (Fig. 4) originated from the variation in responses of biomass (Figure S4) and HI (Figure S5). We attempted to determine whether biomass responses depended on the types of net primary productivity modeling, but there was no strong indication as to which type varied more than the others did (Figure S4). However, simple RUE-type models showed larger variation among models and sites, whereas FvCB-type models showed generally consistent effects of [CO₂] across sites (Figure S4i–p). This is, at a first glance, counterintuitive because FvCB-type models account for the [CO₂] × temperature interaction on leaf photosynthesis; C₃ photosynthesis is enhanced more by elevated [CO₂] at a higher temperature than at a lower temperature (Long, 1991). The RUE approach is generally simple, in some cases assuming that RUE is independent of temperature, but most RUE-type models in this study have a

temperature limitation factor on RUE (Table S2), which generated the site- and temperature-specific response to [CO₂] (Figure S4). Thus, the [CO₂] × temperature interaction was simulated either mechanistically or empirically, but our quantitative understanding of it is still limited. Furthermore, the [CO₂] × temperature interaction on field-grown rice can be complicated by photosynthetic acclimation to elevated [CO₂] (Lin *et al.*, 1997; Adachi *et al.*, 2014), which has not been captured well by most existing crop models (Yin, 2013).

A diverse response of growth duration was observed at +3 °C or higher temperature ranges at all sites, except in Shizukuishi (Figure S3). This was because of the different functions used for temperature response (Zhang & Tao, 2013). Models with an optimum response of development rate to temperature showed a rather prolonged growth duration at high-temperature ranges, whereas models with asymptotic response showed shortened growth duration at almost all temperature ranges. While evidence exists for decreased developmental rate under high temperature above the optimum (Yin *et al.*, 1996), limited temperature range under field experimental conditions makes it difficult to obtain robust parameter estimations, which results in a large variation in predicted growth durations among models (Van Oort *et al.*, 2011). Nevertheless, biomass response to increasing temperatures was rather consistent across models (Figure S4), indicating that increased respiration and/or reduced photosynthesis (e.g., due to increased senescence) at high temperatures overrode a positive effect through prolonged growth duration. Experimental evidence, however, does not necessarily support reduction in biomass by high temperature (Kim *et al.*, 1996a; Ziska *et al.*, 1997; Cheng *et al.*, 2009), suggesting that none of our models account for plant acclimation to high growth temperature. The gap between modeled and observed difference in response to temperature is an important area for model improvement.

Evidence exists that HI is highly sensitive to heat because of increased occurrence of spikelet sterility (Kim *et al.*, 1996b; Prasad *et al.*, 2006). Some models account for the direct impact of heat-induced sterility, but those without the direct effect also showed a decrease in HI with the increase in temperature (Figure S5a–h). They include the differential effects of temperature on sink size and/or source capacity and/or on vegetative and grain-filling periods. These mechanisms also need to be examined in more detailed experiments that determine critical processes affecting HI in response to increasing temperature.

Simulated HI tended to decrease with increasing [CO₂] (Figure S5i–p). In early Free-Air CO₂ Enrichment (FACE) studies, HI was slightly decreased under

elevated [CO₂] (Kim *et al.*, 2003; Yang *et al.*, 2006), but recent FACE studies showed that the effect of [CO₂] on HI was more negative under both low- and high-temperature conditions than under moderate temperature conditions (Shimono *et al.*, 2008; Hasegawa *et al.*, 2013). Some chamber studies reported that elevated [CO₂] increased heat-induced spikelet sterility (Matsui *et al.*, 1997), presumably because of reduced transpirational cooling and increased canopy or panicle temperature under elevated [CO₂] (Yoshimoto *et al.*, 2005). Some models do account for this, but most rice models used in this study do not take canopy or panicle temperature effects into account. In our sensitivity simulation exercise, five out of 12 models predicted a significant negative [CO₂] × temperature interaction on yield, whereas seven models predicted an insignificant negative interaction. Moreover, any [CO₂] × temperature interaction on grain yield can be rice genotype-specific and may be modulated by diurnal temperature difference and other climatic factors such as drought and humidity (Cho & Oki, 2012). We need to increase our field observations of these interactions on subprocesses of yield formation to improve our understanding and quantification on yields in crop models.

A rice model ensemble helps to reduce uncertainty

Based on our results with 13 rice crop models, one could argue whether any single model can be reliably used to predict rice yields under a range of environments. Even after the calibration of each model across four sentinel sites that represent different climatic environments, differences in predicted yields among models, in specific years and sites, or as 30-year averages, were often up to 100% (Figs 1 and 2). Differences in model-predicted yield and experimentally measured yield at our sentinel sites varied from 5% to 60%, depending on the site, management practices, and rice cultivars. No single model consistently provided reliable predictions of yield across sites, management practices, rice cultivars, and years.

Despite considerable variation among models for response to CO₂, temperature and their interaction, we consider the ensemble predictions as an indication of the likely response of rice to those two climate change factors. The ensemble response of all crop models used in the study (mean) showed that predicted response to CO₂ was an increase of 20.6% to 26.2% with a doubling of CO₂ from 360 to 720 μmol mol⁻¹ (Fig. 4i–l, Table S3) with the ensemble differences being associated with warm vs. cool sites. The ensemble response appears to be almost the same as literature-reported responses (Horie *et al.*, 2000; Ainsworth, 2008). Similarly, the crop

model ensemble response (mean) to temperature suggested that rising temperature will reduce yield at all sites, with reductions of 4.8% at the coolest site (Shizukuishi) to 27.3% at the warmest site (Ludhiana) for +3 °C, and 27.9% to 55.2% reduction for +6 °C (Fig. 4a–d, Table 2). The ensemble model response to increased temperature is also close to literature-reported response (Hatfield *et al.*, 2011).

The absolute difference between measured and predicted yields was reduced by using an ensemble of any of the 13 rice models. The use of an ensemble of at least eight models calibrated only for phenology, or five models calibrated in detail against well-controlled field experiments, can result in yield predictions with 15% uncertainty (Fig. 5), which is equivalent to the typical deviation level of measured yield in agronomic field experiments (Belder *et al.*, 2007). Despite inherent differences in the growing environments between rice and wheat crops, the number of models required for robust estimation was strikingly similar to that presented by the AgMIP Wheat team (Asseng *et al.*, 2013) and Maize team (Bassu *et al.*, 2014). Given the current accuracy of rice crop models, we suggest that the use of multiple crop models gives a better yield prediction than a single model across a range of environments. However, reduction in uncertainty under historic climatic conditions does not imply reduced uncertainty for the future. Currently, an effort is underway to test and improve the modeling of [CO₂] and temperature sensitivity against response data by collaborating with previous and ongoing FACE and chamber studies.

Acknowledgements

This study was partly financed by the International Rice Research Institute (IRRI) through the Global Rice Science Partnership (GRiSP). We are grateful to the members of the AgMIP Leadership team for their support and to Senthil Asseng and Frank Ewert for sharing their insights in the AgMIP Wheat team. IRRI (via Dr. Sudhir Yadav and Dr. Roland Buresh), Nanjing Agricultural University (China), and the National Institute for Agro-Environmental Sciences (Japan) kindly provided the experimental data used in this study. We acknowledge the global climate modeling groups, the Program for Climate Model Diagnosis and Intercomparison (PCMDI), and the WCRP's Working Group on Coupled Modelling (WGCM) for making the WCRP CMIP3 multi-model dataset available. Support for this dataset is provided by the Office of Science, U.S. Department of Energy. Dr. Ruane's participation in the study was supported by the NASA Modeling, Analysis, and Prediction Program.

References

- Adachi M, Hasegawa T, Fukayama H *et al.* (2014) Soil and water warming accelerates phenology and down-regulation of leaf photosynthesis of rice plants grown under free-air CO₂ enrichment (FACE). *Plant & Cell Physiology*, **55**, 370–380.

- Ainsworth EA (2008) Rice production in a changing climate: a meta-analysis of responses to elevated carbon dioxide and elevated ozone concentration. *Global Change Biology*, **14**, 1642–1650.
- Asseng S, Ewert F, Rosenzweig C *et al.* (2013) Uncertainty in simulating wheat yields under climate change. *Nature Climate Change*, **3**, 827–832.
- Bassu S, Brisson N, Durand J-L *et al.* (2014) How do various maize crop models vary in their responses to climate change factors? *Global Change Biology*, **20**, 2301–2320.
- Belder P, Bouman BAM, Spiertz JHJ (2007) Exploring options for water saving in low-land rice using a modeling approach. *Agricultural Systems*, **92**, 91–114.
- Cheng W, Sakai H, Yagi K, Hasegawa T (2009) Interactions of elevated [CO₂] and night temperature on rice growth and yield. *Agricultural and Forest Meteorology*, **149**, 51–58.
- Cho J, Oki T (2012) Application of temperature, water stress, CO₂ in rice growth models. *Rice*, **5**, 10.
- De Wit CT (1978) *Simulation of Assimilation, Respiration and Transpiration of Crops*. Pudoc, Wageningen, The Netherlands.
- Easterling WE, Aggarwal PK, Batima P *et al.* (2007) Food, fibre and forest products. In: *Climate Change 2007: Impacts, Adaptation and Vulnerability*. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (eds Parry ML, Canziani OF, Palutikof JP, der van Linden PJ, Hanson CE), pp. 273–313. Cambridge University Press, Cambridge, UK.
- Farquhar GD, von Caemmerer S, Berry JA (1980) A biochemical model of photosynthesis CO₂ assimilation in leaves of C₃ species. *Planta*, **149**, 78–90.
- Godfray HCJ, Beddington JB, Crute IR *et al.* (2010) Food security: the challenge of feeding 9 billion people. *Science*, **327**, 812–818.
- Hasegawa T, Ishimaru T, Kondo M, Kuwagata T, Yoshimoto M, Fukuoka M (2011) Spikelet sterility of rice observed in the record hot summer of 2007 and the factors associated with its variation. *Journal of Agricultural Meteorology*, **67**, 225–232.
- Hasegawa T, Sakai H, Tokida T *et al.* (2013) Rice cultivar responses to elevated CO₂ at two free-air CO₂ enrichment (FACE) sites in Japan. *Functional Plant Biology*, **40**, 148–159.
- Hatfield JL, Boote KJ, Kimball BA *et al.* (2011) Climate impacts on agriculture: implications for crop production. *Agronomy Journal*, **103**, 351–370.
- Horie T, Baker JT, Nakagawa H, Matsui T, Kim HY (2000) Crop ecosystem responses to climatic change: rice. In: *Climate Change and Global Crop Productivity* (eds Reddy KR, Hodges, HF), pp. 81–106. CAB International, New York, USA.
- Jagadish SVK, Craufurd PQ, Wheeler TR (2007) High temperature stress and spikelet fertility in rice (*Oryza sativa* L.). *Journal of Experimental Botany*, **58**, 1627–1635.
- Kim HY, Horie T, Nakagawa H, Wada K (1996a) Effects of elevated CO₂ concentration and high temperature on growth and yield of rice: 1. The effect on development, dry matter production and some growth characteristics. *Japanese Journal of Crop Science*, **65**, 634–643.
- Kim HY, Horie T, Nakagawa H, Wada K (1996b) Effects of elevated CO₂ concentration and high temperature on growth and yield of rice: 2. the effect on yield and its components of Akihikari rice. *Japanese Journal of Crop Science*, **65**, 644–651.
- Kim HY, Loeffering M, Kobayashi K, Okada M, Mitchell MW, Gumpertz M (2003) Effects of free-air CO₂ enrichment and nitrogen supply on the yield of temperate paddy rice crops. *Field Crops Research*, **83**, 261–270.
- Lin W, Ziska LH, Namuco OS, Bai K (1997) The interaction of high temperature and elevated CO₂ on photosynthetic acclimation of single leaves of rice *in situ*. *Physiologia Plantarum*, **99**, 178–184.
- Long SP (1991) Modification of the response of photosynthetic productivity to rising temperature by atmospheric CO₂ concentrations: has its importance been underestimated? *Plant, Cell and Environment*, **14**, 729–739.
- Maclean JL, Dawe D, Hardy B, Hettel GP (2002) *Rice Almanac*, 3rd edn. CABI Publishing, Wallingford, UK. [cited 2012 Nov. 13].
- Matsui T, Namuco OS, Ziska LH, Horie T (1997) Effects of high temperature and CO₂ concentration on spikelet sterility in Indica rice. *Field Crops Research*, **51**, 213–219.
- Meehl GA, Covey C, Taylor KE *et al.* (2007a) THE WCRP CMIP3 Multimodel Dataset: a new era in climate change research. *Bulletin of the American Meteorological Society*, **88**, 1383–1394.
- Meehl GA, Stocker TF, Collins WD *et al.* (2007b) Global climate projections. In: *Climate Change 2007: The Physical Science Basis*. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (eds Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL), pp. 747–845. Cambridge University Press, Cambridge, UK.
- Monteith JL, Moss CJ (1977) Climate and the efficiency of crop production in Britain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **281**, 277–294.
- Prasad PVV, Boote KJ, Allen LH, Sheehy JE, Thomas JMG (2006) Species, ecotype and cultivar differences in spikelet fertility and harvest index of rice in response to high temperature stress. *Field Crops Research*, **95**, 398–411.
- Rosenzweig C, Jones JW, Hatfield JL *et al.* (2013) The Agricultural Model Intercomparison and Improvement Project (AgMIP): protocols and pilot studies. *Agricultural and Forest Meteorology*, **170**, 166–182.
- Rosenzweig C, Elliott J, Deryng D *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 of the United States of America*, **111**, 3268–3273.
- Satake T (1976) Sterile-type cool injury in paddy rice plants. *Climate and Rice*, pp. 281–300. International Rice Research Institute, Los Baños, Philippines.
- Satake T, Yoshida S (1978) High temperature-induced sterility in indica rices at flowering. *Proceedings of the Crop Science Society of Japan*, **47**, 6–17.
- Shimono H, Okada M, Yamakawa Y, Nakamura H, Kobayashi K, Hasegawa T (2008) Rice yield enhancement by elevated CO₂ is reduced in cool weather. *Global Change Biology*, **14**, 276–284.
- Sudhir-Yadav, Gill G, Humphreys E *et al.* (2011a) Effect of water management on dry seeded and puddled transplanted rice Part 1: Crop performance. *Field Crops Research*, **120**, 112–122.
- Sudhir-Yadav, Li T, Humphreys E *et al.* (2011b) Evaluation and application of ORYZA2000 for irrigation scheduling of puddled transplanted rice in North West India. *Field Crops Research*, **122**, 104–117.
- Tao F, Hayashi Y, Zhang Z, Sakamoto T, Yokozawa M (2008) Global warming, rice production and water use in China: developing a probabilistic assessment. *Agricultural and Forest Meteorology*, **148**, 94–110.
- Van Oort PAJ, Zhang T, de Vries ME, Heinemann AB, Meinke H (2011) Correlation between temperature and phenology prediction error in rice (*Oryza sativa* L.). *Agricultural and Forest Meteorology*, **151**, 1545–1555.
- Wassmann R, Jagadish SVK, Sumfleth K, Pathak H, Singh RK, Heuer S (2009) Regional vulnerability of climate change impacts on Asian rice production and scope for adaptation. *Advances in Agronomy*, **102**, 91–133.
- Yang L, Huang J, Yang H *et al.* (2006) The impact of free-air CO₂ enrichment (FACE) and N supply on yield formation of rice crops with large panicle. *Field Crops Research*, **98**, 141–150.
- Yin X (2013) Improving ecophysiological simulation models to predict the impact of elevated atmospheric CO₂ concentration on crop productivity. *Annals of Botany*, **112**, 465–475.
- Yin X, Kropff MJ, Goudriaan J (1996) Differential effects of day and night temperature on development to flowering in rice. *Annals of Botany*, **77**, 203–213.
- Yoshimoto M, Oue H, Takahashi N (2005) The effects of FACE (Free-Air CO₂ Enrichment) on temperatures and transpiration of rice panicles at flowering stage. *Journal of Agricultural Meteorology*, **60**, 597–600.
- Zhang S, Tao F (2013) Modeling the response of rice phenology to climate change and variability in different climatic zones: comparisons of five models. *European Journal of Agronomy*, **45**, 165–176.
- Ziska LH, Namuco O, Moya T, Quilang J (1997) Growth and yield response of field-grown tropical rice to increasing carbon dioxide and air temperature. *Agronomy Journal*, **89**, 45–53.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Total above-ground biomass and grain yield measured and simulated by different models for each treatment in four sentinel datasets. Acronyms for the treatments are given in Table 1, but they generally start with year and N or irrigation treatments. Values followed by N are N application rates in kg ha^{-1} .

Figure S2. The historical yields for regions of four sentinel datasets from 1980 to 2010 (dots), their polynomial regression presented lines (thick lines), 95% confidence ranges of polynomial functions defined by two thin lines (a to d), and the time-series residuals between recorded yields and those predicted by polynomial functions (e to h).

Figure S3. The response of modeled total growth duration to temperature increases (ΔT) under two atmospheric $[\text{CO}_2]$ conditions: $360 \mu\text{mol mol}^{-1}$ (a to d) and $720 \mu\text{mol mol}^{-1}$ (e to f), respectively.

Figure S4. Response of modeled total above-ground biomass to change in air temperature and atmospheric $[\text{CO}_2]$.

Figure S5. Response of modeled harvest index to change in air temperature and atmospheric $[\text{CO}_2]$.

Table S1. Model name, institutional affiliation, and reference for a detailed explanation of the models included in our study.

Table S2. Forcing variables and target processes of each model.

Table S3. The list of the parameters that were subject to changes in the model calibration phase, and the comparison between their values used in blind simulation (Blind) and values after the calibration (Calibrated) for each sentinel site.

Table S4. Analysis of simulation results for individual models over all treatments in four sentinel datasets.