Chapter 11

Statistical Analysis of Large Simulated Yield Datasets for Studying Climate Effects

David Makowski[1], Senthold Asseng[2], Frank Ewert[3], Simona Bassu[1], Jean-Louis Durand[4], Pierre Martre[5,6], Myriam Adam[7], Pramod K. Aggarwal[8], Carlos Angulo[3], Christian Baron[9], Bruno Basso[10], Patrick Bertuzzi[11], Christian Biernath[12], Hendrik Boogaard[13], Kenneth J. Boote[14], Nadine Brisson[1,11,#], Davide Cammarano[2], Andrew J. Challinor[15,16], Sjakk J. G. Conijn[17], Marc Corbeels[45], Delphine Deryng[19], Giacomo De Sanctis[11], Jordi Doltra[20], Sebastian Gayler[21], Richard Goldberg[22], Patricio Grassini[23], Jerry L. Hatfield[24], Lee Heng[25], Steven Hoek[13], Josh Hooker[26], Tony L. A. Hunt[27], Joachim Ingwersen[28], Cesar Izaurralde[29], Raymond E. E. Jongschaap[17], James W. Jones[2], Armen R. Kemanian[30], Christian Kersebaum[31], Soo-Hyung Kim[32], Jon Lizaso[33], Christoph Müller[34], Naresh S. Kumar[35], Claas Nendel[31], Garry J. O'Leary[36], Jorgen E. Olesen[37], Tom M. Osborne[38], Taru Palosuo[18], Maria V. Pravia[30,39], Eckart Priesack[12], Dominique Ripoche[11], Cynthia Rosenzweig[22], Alexander C. Ruane[22], Fredirico Sau[40], Mickhail A. Semenov[41], Iurii Shcherbak[10], Pasquale Steduto[42], Claudio Stöckle[43], Pierre Stratonovitch[41], Thilo Streck[28], Iwan Supit[44], Fulu Tao[45], Edmar I. Teixeira[46], Peter Thorburn[47], Denis Timlin[48], Maria Travasso[49], Reimund Rötter[18], Katharina Waha[34], Daniel Wallach[50], Jeffrey W. White[51], Jimmy R. Williams[52], and Joost Wolf[44]

[1]*L'Institut national de la recherche agronomique, Thiverval-Grignon, France*

[2]*University of Florida, Gainesville, FL, USA*

[3]*University of Bonn, Bonn, Germany*

[4]*Unité de recherche pluridisciplinaire sur la prairie et les plantes fourragères, Lusignan, France*

[5]*Genetic, Diversity and Ecophysiology of Cereals (GDEC), Clermont-Ferrand, France*

[6]*Blaise Pascal University, Aubière, France*

[7]*CIRAD, Montpellier, France*

[8]*CCAFS, New Delhi, India*

[9]*CIRAD, Montpellier, France*

[10]*Michigan State University, East Lansing, MI, USA*

[11] *L'Institut national de la recherche agronomique, Avignon, France*
[12] *German Research Center for Environmental Health, Neuherberg, Germany*
[13] *Centre for Geo-Information, Alterra, the Netherlands*
[14] *University of Florida, Gainesville, FL, USA*
[15] *University of Leeds, Leeds, UK*
[16] *International Centre for Tropical Agriculture (CIAT), Cali, Colombia*
[17] *Wageningen University and Research Centre, Wageningen, the Netherlands*
[18] *MTT Agrifood Research Finland, Mikkeli, Finland*
[19] *University of East Anglia, Norwich, UK*
[20] *Cantabrian Agricultural Research and Training Centre (CIFA), Muriedas, Spain*
[21] *University of Tübingen, Tübingen, Germany*
[22] *NASA Goddard Institute for Space Studies, New York, NY, USA*
[23] *University of Nebraska-Lincoln, Lincoln, USA*
[24] *National Laboratory for Agriculture and Environment, Ames, IA, USA*
[25] *International Atomic Energy Agency (IAEA), Vienna, Austria*
[26] *University of Reading, Reading, UK*
[27] *University of Guelph, Guelph, Ontario, Canada*
[28] *University of Hohenheim, Stuttgart, Germany*
[29] *Joint Global Change Research Institute, College Park, MD, USA*
[30] *Instituto Nacional de Investigación Agropecuaria (INIA), Treinta y Tres, Uruguay*
[31] *Leibniz Centre for Agricultural Landscape Research, Müncheberg, Germany*
[32] *University of Washington, Seattle, WA, USA*
[33] *Univ. Politécnica of Madrid, Madrid, Spain*
[34] *Potsdam Institute for Climate Impact Research, Potsdam, Germany*
[35] *Indian Agricultural Research Institute, New Delhi, India*
[36] *Landscape & Water Sciences, Department of Environment and Primary Industries, Horsham, Australia*
[37] *Aarhus University, Tjele, Denmark*
[38] *University of Reading, Reading, UK*
[39] *The Pennsylvania State University, University Park, PA, USA*
[40] *Univ. Politécnica of Madrid, Madrid, Spain*
[41] *Rothamsted Research, Harpenden, UK*
[42] *Food and Agriculture Organization of the UN (FAO), Rome, Italy*
[43] *Washington State University, Pullman, WA, USA*
[44] *Wageningen University, Wageningen, the Netherlands*

[45] *Chinese Academy of Science, Beijing, China*

[46] *The New Zealand Institute for Plant & Food Research Limited, Canterbury, New Zealand*

[47] *Commonwealth Scientific and Industrial Research Organisation (CSIRO) Ecosystem Sciences, Queensland, Australia*

[48] *United States Department of Agriculture, Agricultural Research Service (USDA-ARS), Beltsville, MD, USA*

[49] *Institute for Climate and Water, Castelar, Argentina*

[50] *Agrosystèmes et développement territorial (AGIR), Castanet-Tolosan, France*

[51] *Arid-Land Agricultural Research Center, Maricopa, AZ, USA*

[52] *Texas A&M University, TX, USA*

[#] *Dr. Nadine Brisson passed away in 2011 while this work was being carried out.*

Introduction

Many studies have been carried out during the last decade to study the effect of climate change on crop yields and other key crop characteristics. In these studies, one or several crop models were used to simulate crop growth and development for different climate scenarios that correspond to different projections of atmospheric CO_2 concentration, temperature, and rainfall changes (Semenov *et al.*, 1996; Tubiello and Ewert, 2002; White *et al.*, 2011). The Agricultural Model Intercomparison and Improvement Project (AgMIP; Rosenzweig *et al.*, 2013) builds on these studies with the goal of using an ensemble of multiple crop models in order to assess effects of climate change scenarios for several crops in contrasting environments.

These studies generate large datasets, including thousands of simulated crop yield data. They include series of yield values obtained by combining several crop models with different climate scenarios that are defined by several climatic variables (temperature, CO_2, rainfall, etc.). Such datasets potentially provide useful information on the possible effects of different climate change scenarios on crop yields. However, it is sometimes difficult to analyze these datasets and to summarize them in a useful way due to their structural complexity; simulated yield data can differ among contrasting climate scenarios, sites, and crop models. Another issue is that it is not straightforward to extrapolate the results obtained for the scenarios to alternative climate change scenarios not initially included in the simulation protocols. Additional dynamic crop model simulations for new climate change scenarios are an option but this approach is costly, especially when a large number of crop models are used to generate the simulated data, as in AgMIP.

Statistical models have been used to analyze responses of measured yield data to climate variables in past studies (Lobell *et al.*, 2011), but the use of a statistical model to analyze yields simulated by complex process-based crop models is a rather new idea. We demonstrate herewith that statistical methods can play an important role in analyzing simulated yield data sets obtained from the ensembles of process-based crop models. Formal statistical analysis is helpful to estimate the effects of different climatic variables on yield, and to describe the between-model variability of these effects.

These statistical methods can also be used to develop meta-models, i.e., statistical models that summarize process-based crop models, enabling scientists to explore the effects of new climate change scenarios. This approach is illustrated with two simulated yield datasets obtained by AgMIP for maize and wheat that were generated by using ensembles of process-based crop models. The yield datasets were used to develop a meta-model that provides a simplified representation of the original ensemble of crop models. The proposed meta-model is a statistical regression with random coefficients that describe the variability of the simulated yield data across the original crop models. Once the statistical regression has been fitted to the simulated yield datasets, the meta-model can predict the effects of climate changes among the crop models and can thus be used to study alternative climate change scenarios that were not initially simulated. Finally, the results obtained with the meta-model were used to assess the risk of yield loss for maize and wheat as functions of temperature and CO_2 concentration change.

Materials and Methods

Simulated yield data

Yield data were simulated with 19 maize crop models and 26 wheat crop models in four contrasting sites for each species located in France (Lusignan), USA (Ames), Brazil (Rio Verde), and Tanzania (Morogoro) for maize, and in the Netherlands (Wageningen), Argentina (Balcarce), India (New Delhi), and Australia (Wongan Hills) for wheat.

Several climate scenarios were considered. For both species, four temperature changes (+0, +3, +6, +9°C) and five atmospheric CO_2 concentration changes (+0, +90, +180, +270, +360 ppm; where the baseline CO_2 concentration was set at 360 ppm) were considered over the 1980–2010 time-period. In addition, two precipitation changes (0% and −30%) were considered for maize and three rates of nitrogen fertilization (50%, 100%, and 150% of reference rates) were considered for wheat. The other scenarios were defined by combining temperature, CO_2 concentration, and precipitation changes. Thirty years of yield data were generated with each crop model for each of the climate scenarios (Figs. 1 and 2) and the simulated yield values were averaged over the years. Details of the maize and wheat protocols can be found in Bassu *et al.* (2014) and Asseng *et al.* (2013), respectively.

Statistical model

Simulated maize and wheat yield data were analyzed using two-level statistical random-effect models (Pinheiro and Bates, 2000). The following statistical model was used to analyze the maize yield data for each site separately:

Level 1, within crop model

$$Y_{ij} = \alpha_{0i} + \alpha_{1i}\Delta T_{ij} + \alpha_{2i}\Delta T_{ij}^2 + \alpha_{3i}\Delta C_{ij} + \alpha_{4i}\Delta C_{ij}^2 + \alpha_{5i}\Delta C_{ij}\Delta T_{ij} + \alpha_{6i}\Delta R_{ij} + \varepsilon_{ij} \quad (1)$$

where Y_{ij} is the 30-year mean yield simulated for the ith crop model (where $i = 1, \dots, P$) and for the jth scenario (where $j = 1, \dots, Q_i$). The variables ΔT_{ij}, ΔC_{ij},

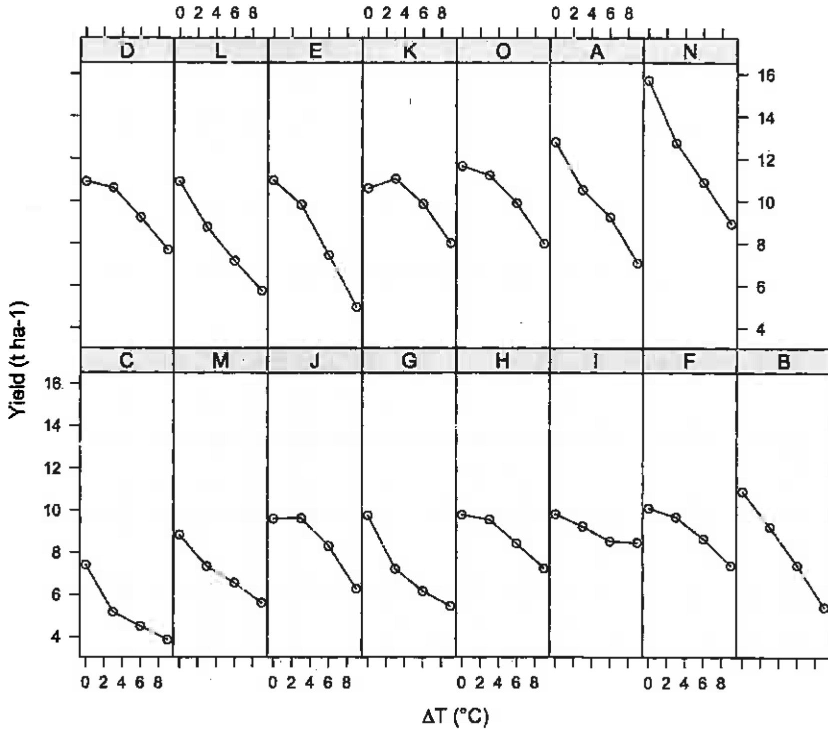


Fig. 1. Maize yields simulated by 15 crop models in Lusignan, France for different levels of temperature change (each letter corresponds to one crop model). The reported yield data were obtained for a change of CO_2 concentration equal to +180 ppm and for the baseline (1981–2010) level of precipitation. Only 15 out of the 19 available crop models simulated this level of [CO_2] and were thus included in the figure.

and ΔR_{ij} are the respective changes from the 1980–2010 baseline of temperature, atmospheric CO_2 concentration, and precipitation for model i and scenario j. The value of ε_{ij} is assumed to be independently and identically distributed such that $\varepsilon_{ij} \sim N(0, \tau^2)$ where τ^2 is the variance describing the residual error. The coefficients, α_{ki} (where $k = 0, \ldots, 6$), are seven random regression coefficients distributed according to independent Gaussian probability distributions.

Level 2, between crop models

$$\alpha_{ki} \sim N(\mu_k, \sigma_k^2), \quad k = 0, \ldots, 6 \tag{2}$$

The distribution parameters, μ_k (where $k = 0, \ldots, 6$), are the seven mean regression coefficient values that represent the mean yield baseline (μ_0) and mean effects of temperature, CO_2, temperature–CO_2 interaction, and rainfall ($\mu_1, \ldots, \mu_6$) over the P crop models. The seven variances, σ_k^2 (where $k = 0, \ldots, 6$), describe the between-model variability of the random regression coefficients.

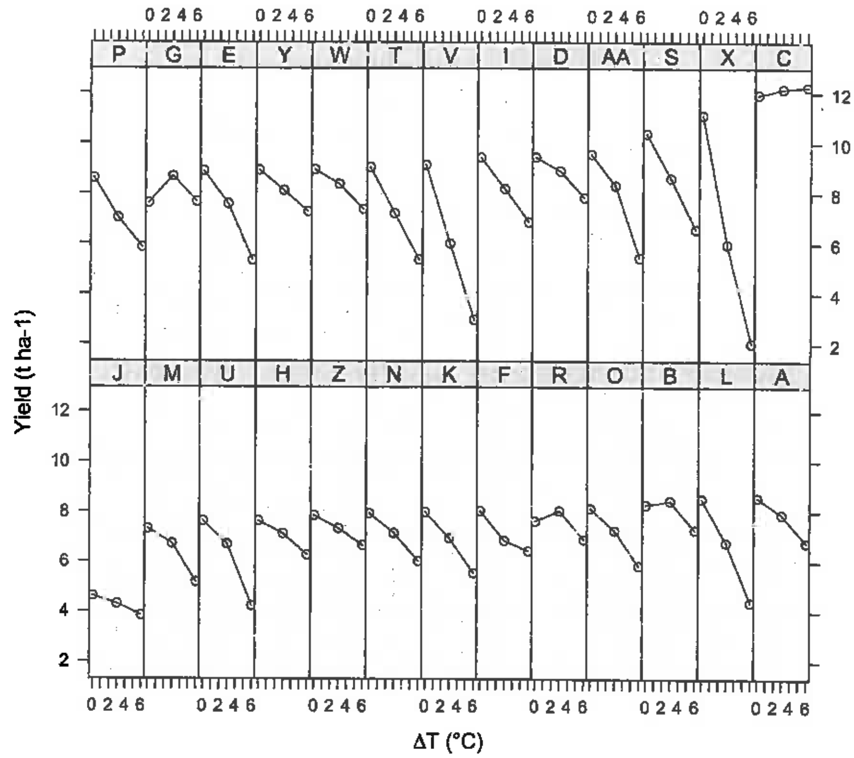


Fig. 2. Wheat yields simulated by the 26 crop models in Wageningen, the Netherlands, for different levels of temperature change (each letter corresponds to one crop model). The reported yield data were obtained for a change of CO_2 concentration equal to +180 ppm and for the baseline (1981–2010) level of fertilization.

This statistical model assumes that the ensemble of P crop models is a sample taken within a population that includes all possible crop models for a given crop while flexibly allowing for the incorporation of additional crop models in the future. The probability distributions defined by Equation (2) describe the between-crop-model variability of the yield response to climate change within the whole population of crop models. These probability distributions cover the ranges of climate effects considered in different crop models. The relationship defined by Equation (1) is assumed to be valid for all crop models, but its parameters, α_{ki}, are assumed to vary among crop models. However, this statistical model only describes 30-year mean yield responses and is not intended to describe the year-to-year variability of crop yields. Creating a statistical model that captured year-to-year variability would have required extra random terms and additional parameters, which would have overly complicated the calculated model and thus was not considered here.

The same form of model was used for wheat as was used for maize, but replaced the rainfall variable, ΔR_{ij}, with a variable that describes the level of N fertilization in Equation (1). For the wheat pilot studies, precipitation was held constant in the wheat simulation protocol.

The population parameters of the statistical model μ_k, σ_k^2, and τ^2 were estimated by restricted maximum likelihood. The model-specific regression coefficients, α_{ki} (where $k = 0, \ldots, 6$ and $i = 1, \ldots, P$), were estimated by using the best linear unbiased predictor using the R software package "nlme" (Pinheiro and Bates, 2000) and the estimated values will be henceforth referred to as α_{eki}. The model was fitted for each crop and each site separately, but was calculated for all crop models together. Results were analyzed site by site.

Assessment of the statistical model

The statistical model was compared to other statistical models, including models with fewer explanatory variables, models with fewer random coefficients, and a model that includes no random coefficients (i.e., classical linear regression). All the models were compared by using the Akaïke information criterion (AIC; Pinheiro and Bates, 2000), where a lower AIC value corresponds with the better model, and we found that the model defined by Equations (1) and (2) led to lower AICs than the simpler models. The assumption that the residual errors, ε_{ij}, were independent was assessed by developing another statistical model that incorporated correlated residual errors. The AIC value of this model was higher and the estimated correlation coefficients were very low (7.6×10^{-4} and -4.7×10^{-3} for maize and wheat, respectively). The statistical model's quality of fit was also assessed by using graphical analysis. Evaluating the residuals did not reveal any obvious biases due to the assumptions made on the relationships between yields and input factors.

Estimation of the effect of climate change on yield

The statistical model described above was used to compute three different types of outputs:

- The average yield loss/gain due to climate change over the ensemble of crop models.
- The yield gain/loss estimated for individual crop models due to changes in climate variables.
- The probability that the yield loss/gain will exceed a given threshold.

For maize, the average yield difference obtained between a given climate change scenario (characterized by ΔT, ΔC, and ΔR) and the baseline scenario was expressed as:

$$\Delta Y = \mu_1 \Delta T + \mu_2 \Delta T^2 + \mu_3 \Delta C + \mu_4 \Delta C^2 + \mu_5 \Delta C \Delta T + \mu_6 \Delta R \qquad (3)$$

The yield difference described in Equation (3) is averaged over all crop models; this difference corresponds to an average yield gain or to an average yield loss over the 19 submitted crop models. Equation (3) defines a meta-model that simulates the average output of the original ensemble of crop models. This meta-model enables the computation of the yield differences for any changes in temperature, CO_2, and precipitation, ΔT, ΔC, and ΔR, at each of the four considered sites.

For a given crop model i, the yield difference was expressed as:

$$\Delta Y_i = \alpha_{e1i}\Delta T + \alpha_{e2i}\Delta T^2 + \alpha_{e3i}\Delta C + \alpha_{e4i}\Delta C^2 + \alpha_{e5i}\Delta C\Delta T + \alpha_{e6i}\Delta R \quad (4)$$

Equation (4) defines a meta-model simulating the output of the ith crop model. The yield difference (4) is crop model-specific. It corresponds to the climate change effect on yield that would have been obtained with the ith crop model if this crop model were run for a climate change scenario that were characterized by ΔT, ΔC, and ΔR.

The statistical model defined by Equations (1) and (2) was also used to compute the probability that the yield difference will exceed a threshold and, more specifically, the probability of yield gain or of yield loss that results from a change in the temperature, CO_2 concentration, and/or precipitation. This probability was computed from the following Gaussian probability distribution $N(\mu_{\Delta Y}, \sigma^2_{\Delta Y})$ as:

$$\mu_{\Delta Y} = \mu_1\Delta T + \mu_2\Delta T^2 + \mu_3\Delta C + \mu_4\Delta C^2 + \mu_5\Delta C\Delta T + \mu_6\Delta R \quad (5)$$

$$\sigma^2_{\Delta Y} = \sigma_1^2\Delta T^2 + \sigma_2^2\Delta T^4 + \sigma_3^2\Delta C^2 + \sigma_4^2\Delta C^4 + \sigma_5^2\Delta C^2\Delta T^2 + \sigma_6^2\Delta R^2 \quad (6)$$

Similar relationships were derived for wheat by replacing precipitation change with fertilization change in Equations (3) to (6). Note that the variance defined by Equation (6) is not constant but varies as a function of the climate scenario characteristics.

Results and Discussion

Increases in temperature

Figure 3 shows the statistical meta-models' quality of fit calculated for maize and wheat (one site per crop). The distributions of the residuals for both maize and wheat are symmetric and do not display trends. The fitted values obtained with the statistical meta-models are in most cases close to the yields simulated by the processed-based crop models. The residual standard error of the fitted statistical models is equal to 0.19 and 0.22 t/ha for maize and wheat, respectively, in the two sites considered in Fig. 3. A similar quality of fit was obtained for the other sites.

Figure 4 shows the change in the mean yield from the baseline for one maize site (Fig. 4a) and one wheat site (Fig. 4b) as affected by a CO_2 concentration increase

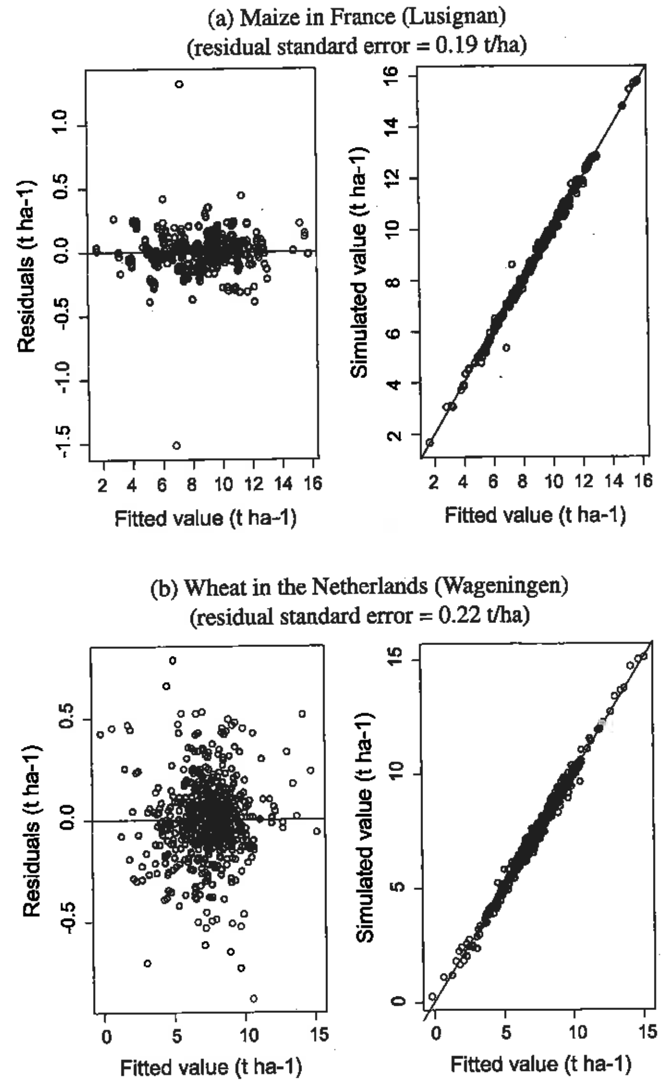


Fig. 3. Residuals of the statistical model and simulated versus fitted values obtained for (a) maize, and (b) wheat (one site per crop, all scenarios). The graphics compare yield values simulated by the crop models and yield values fitted by the statistical model for all scenarios.

of 180 ppm ($[CO_2] = 540$ ppm) and an increase in mean seasonal temperature ranging from 0°C to 6°C. Each emulated model yield as calculated by using the crop model-specific coefficients α_{eki} (where $k = 0, \ldots, 6$ and $i = 1, \ldots, P$) is plotted with a black line and thus can be seen as a proxy of a given crop model. Negative yield differences can be interpreted as mean yield losses and, conversely, positive yield differences can be interpreted as mean yield gains without the need

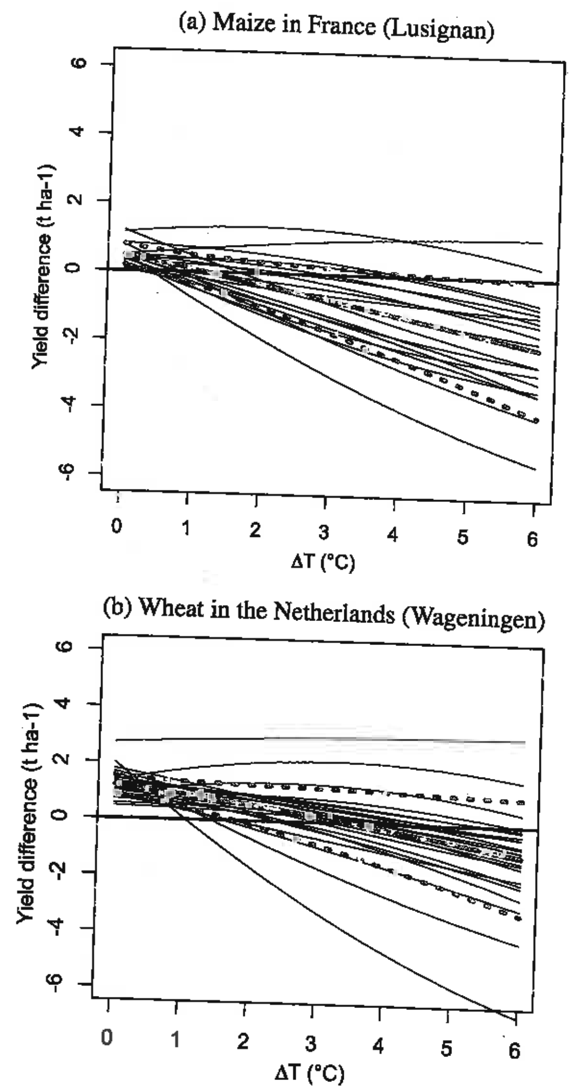


Fig. 4. The effect of climate change on yield for a CO_2 concentration increase of 180 ppm, and for mean temperature increases ranging from 0 to 6°C. Results are reported for (a) maize, and (b) wheat for one site per crop. Yield difference = yield obtained with a climate change scenario − yield obtained with the baseline (1981–2010). Each black curve corresponds to a given crop model. Black curves were obtained with the statistical model by using the crop model-specific coefficients α_{eki} (where $k = 0, \ldots, 6$ and $i = 1, \ldots, P$). The continuous red curve indicates the mean yield response as compared to the baseline scenario, i.e., the effect averaged over all crop models. The red dashed curves indicate the first and third quartiles of climate change effect computed over all crop models. The blue line indicates zero yield difference.

for re-running the original, process-based crop model. The continuous red curve indicates the mean yield response to the given climate scenario as compared to the baseline, i.e., the effect averaged over all crop models. The red dashed curves indicate the first (probability = 0.25) and third (probability = 0.75) quartiles of the climate change effect.

According to Fig. 4, most of the crop models estimate that a temperature increase negatively impacts yield projections of both maize and wheat in all locations. But this effect is highly variable among crop models, with some models predicting little response to temperature. For maize, Fig. 4a illustrates how, on average across the ensemble of crop models, the statistical meta-model emulates a yield loss when temperatures exceed +1.5°C with a CO_2 concentration increase of +180 ppm in Lusignan, France. Contrastingly, the models suggest that wheat is more resilient to heat stress and would require a temperature threshold equal to +3.6°C before experiencing yield losses below the baseline in Wageningen, the Netherlands (Fig. 4b). Based on the emulated projections seen in Fig. 4b, the models indicate that, should climate change be limited to a +3°C change and include a CO_2 concentration increase of +180 ppm, moderate temperature increases could, in fact, lead to gains in wheat production in this location.

Figure 4 also demonstrates the large variability between crop models and displays how this variability increases as a function of temperature. The differences between the first and third quartiles are much larger for higher temperature changes, at the given CO_2 concentration. This result indicates that the differences between crop models are much larger for high-temperature changes than for low-temperature changes. The level of divergence between the predictions of the crop models thus tends to increase with the temperature value considered in the climate change scenarios.

Increases in atmospheric CO_2 concentration

Figure 5 shows the effect of climate change at one site for both maize and wheat yields under increasing levels of CO_2 concentration, ranging from 0 to +360 ppm from the simulated baseline concentration (i.e., 360 to 720 ppm) at a constant temperature increase of +2°C. Figure 5a illustrates how all crop model simulations predict a positive effect on maize yields in Lusignan, France when exposed to increased CO_2 concentrations, but this effect is small and varies among crop models. The mean curve suggests that the effect of a temperature increase of +2°C is negative for the range of CO_2 concentration increases, thus the benefit of the increased CO_2 concentration does not outweigh the negative effect of the temperature increase. On average, the crop models predict a yield loss for maize, even at high CO_2 concentrations, due to this +2°C temperature change.

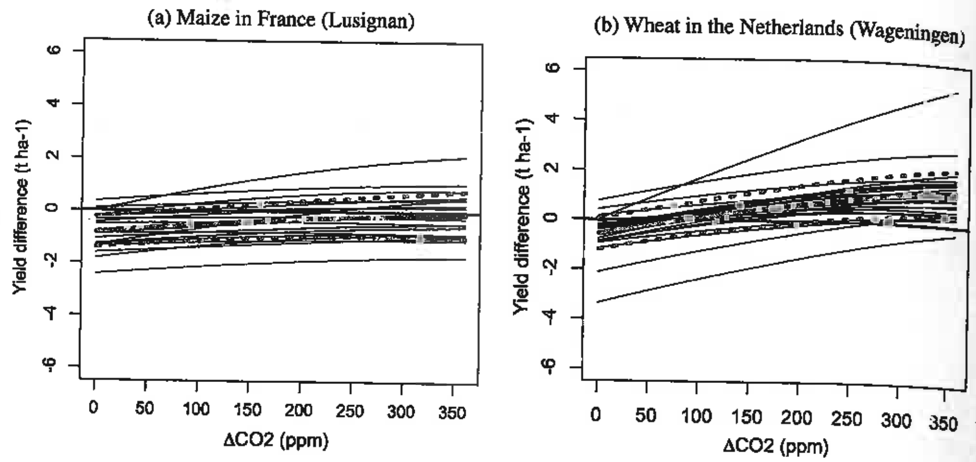
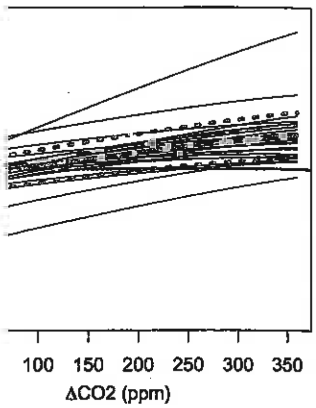


Fig. 5. The effect of climate change on yield for a temperature increase of +2°C, and for levels of CO_2 concentration increase ranging from 0 to +350 ppm. Results are reported for (a) maize and (b) wheat for one site per crop. Yield difference = yield obtained with a climate change scenario — yield obtained with the baseline (1981–2010). Each black curve corresponds to a given crop model. The black curves were obtained with the statistical model by using the crop model-specific coefficients α_{eki} (where $k = 0, \ldots, 6$ and $i = 1, \ldots, P$). The continuous red curve indicates the yield response as compared to the baseline scenario, i.e., the effect averaged over all crop models. The red dashed curves indicate the first and third quartiles of climate change effect computed over all crop models. The blue line indicates zero yield difference.

The fitted curves obtained for CO_2 response of wheat in Wageningen, the Netherlands are displayed in Figure 5b. Compared to maize, the effect of a CO_2 concentration increase is stronger for wheat. This effect is highly variable among crop models; some models demonstrate positive slopes over the range of CO_2 concentrations, whereas others show slopes close to zero. When averaged across all of the crop models, the effect of the +2°C temperature increase is positive (yield gain) as soon as the CO_2 concentration increase reaches 100 ppm.

Probabilistic thresholds

An important advantage of the meta-model is that it handles the interpolation between temperature levels and between CO_2 concentration levels. Our meta-model can thus be used to calculate temperature and [CO_2] thresholds that lead to yield loss or yield gain. Table 1 shows the thresholds of [CO_2] increase required to obtain a probability of maize and wheat yield gain higher than 0.5 (i.e., 50% chance of yield gain). These thresholds were computed for two values of temperature increase (+2 and +4°C) and four sites per crop. It would have been very difficult to compute them without a meta-model. Results show that, for wheat, climate change has 50% chance of resulting in a yield gain if [CO_2] increases by at least +117 ppm (depending on ıt in the Netherlands (Wageningen)



ɾease of +2°C, and for levels of
ɾe reported for (a) maize and (b)
:limate change scenario — yield
ınds to a given crop model. The
crop model-specific coefficients
ırve indicates the yield response
ıll crop models. The red dashed
computed over all crop models.

vheat in Wageningen, the
naize, the effect of a CO_2
t is highly variable among
ver the range of CO_2 con-
'hen averaged across all of
ɜase is positive (yield gain)
pm.

handles the interpolation
ion levels. Our meta-model
holds that lead to yield loss
crease required to obtain a
5 (i.e., 50% chance of yield
f temperature increase (+2
y difficult to compute them
ate change has 50% chance
+117 ppm (depending on

Table 1. Levels of increased CO_2 concentration (ppm) required to obtain a probability of maize and wheat yield gain higher than 0.5 (i.e., 50% chance of yield gain). Concentrations were computed for two values of temperature increase (+2 and +4°C) and four sites per crop. Baseline scenario corresponds to $[CO_2] = 360$ ppm.

	Temperature change	
Site	+2°C	+4°C
Maize	$[CO_2]$	$[CO_2]$
Rio Verde, Brazil	> +360	> +360
Lusignan, France	> +360	> +360
Morogoro, Tanzania	> +360	> +360
Ames, USA	+269	> +360
Wheat	$[CO_2]$	$[CO_2]$
Balcarce, Argentina	+117	+358
Wongan Hills, Australia	0	+59
New Delhi, India	+112	+278
Wageningen, the Netherlands	+83	+222

the site) and if temperature concurrently increases by +2°C (Table 1). Required levels of $[CO_2]$ increases are much higher for maize (Table 1). Our meta-model has another advantage, although not fully illustrated here; it can be used to quantify the effects of temperature, $[CO_2]$, and precipitation on yields, their interactions, and their variability between the four considered sites and between all the considered crop models. It thus constitutes a powerful tool for exploring complex crop model responses to climate factors.

Conclusions

Our study shows that yield data simulated by an ensemble of complex dynamic-process crop models can be summarized by using statistical meta-models that are based on random coefficient regressions. These statistical models describe the between-crop model variability of the simulated yield data by using probability distributions. They can be used to compute key quantities such as mean yield loss, percentiles of yield loss, and probabilities of yield loss as functions of temperature change and CO_2 concentration change. These statistical meta-models can thus be helpful for analyzing the risk of yield loss due to climate change at the locations where the original simulations were conducted. The capabilities of the meta-models were illustrated by using two yield datasets generated by the

AgMIP pilot model intercomparison studies for maize and wheat. The results showed that:

- High CO_2 concentrations are likely to outweigh the negative effect of increasing temperature, leading to yield gains for wheat with climate change, but not for maize.
- Thresholds of temperature increase and CO_2 concentrations that lead to a specific risk level of maize and wheat yield loss can be computed by using the meta-model.
- The divergence between the maize and wheat crop models and therefore the uncertainty in the simulated results increases as a function of temperature (the higher the temperature change, the higher the between-crop model variability).

In the future, the proposed meta-model could be extended in two different ways. First, it could be applied to a dataset that includes simulations obtained for a higher number of sites. It will then be possible to include co-variables that describe site characteristics (e.g., soil type, agricultural practices) in the meta-model in order to explain the origin of the between-site variability. Second, our meta-model could be extended in order to describe the between-year variability of yields and to analyze the risk of extreme yield values for different climate scenarios. This could be achieved by including one or several additional random effects in the meta-model and by fitting this model to yearly yield data.

Acknowledgements

G. J. O'Leary was supported by the Victorian Department of Environment and Primary Industries and the Australian Grains Research and Development Corporation. P. Bertuzzi, J.-L. Durand, D. Makowski, P. Marte, D. Ripoche, and D. Wallach were partly supported by the INRA ACCAF meta-program. S. G. was supported by a grant from the Ministry of Science, Research and Arts of Baden-Württemberg (AZ Zu33-721.3-2) and the Helmholtz Centre for Environmental Research — UFZ, Leipzig.

References

Asseng, S., Ewert, F., Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., Boote, K. J., Thorburn, P. J., Rotter, R. P., Cammarano, D., Brisson, N., Basso, B., Martre, P., Aggarwal, P. K., Angulo, C., Bertuzzi, P., Biernath, C., Challinor, A. J., Doltra, J., Gayler, S., Goldberg, R., Grant, R., Heng, L., Hooker, J., Hunt, L. A., Ingwersen, J., Izaurralde, R. C., Kersebaum, K. C., Muller, C., Naresh Kumar, S., Nendel, C., O'Leary, G., Olesen, J. E., Osborne, T. M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M. A., Shcherbak, I., Steduto, P., Stockle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., Wallach, D., White, J. W., Williams, J. R., and Wolf,

J. (2013). Uncertainty in simulating wheat yields under climate change, *Nat. Clim. Change*, **3**, 827–832.

Bassu, S., Brisson, N., Durand, J.-L., Boote, K., Lizaso, J., Jones, J. W., Rosenzweig, C., Ruane, A. C., Adam, M., Baron, C., Basso, B., Biernath, C., Boogaard, H., Conijn, S., Corbeels, M., Deryng, D., De Sanctis, G., Gayler, S., Grassini, P., Hatfield, J., Hoek, S., Izaurralde, C., Jongschaap, R., Kemanian, A. R., Kersebaum, K. C., Kim, S.-H., Kumar, N. S., Makowski, D., Müller, C., Nendel, C., Priesack, E., Pravia, M. V., Sau, F., Shcherbak, I., Tao, F., Teixeira, E., Timlin, D., and Waha, K. (2014). How do various maize crop models vary in their responses to climate change factors?, *Global Change Biol.*, **20**(7), 2301–2320.

Lobell, D. B., Schlenker, W., and Costa-Roberts, J. (2011). Climate trends and global crop production since 1980, *Science*, **333**, 616–620.

Pinheiro, J. C., and Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*, Springer, New York.

Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., Boote, K. J., Thorburn, P., Antle, J. M., Nelson, G. C., Porter, C., Janssen, S., Asseng, S., Basso, B., Ewert, F., Wallach, D., Baigorria, G., and Winter, J. M. (2013). The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies, *Agric. Forest Meteorol.*, **170**,166–182.

Semenov, M. A., Wolf, J., Evans, L. G., Eckersten, H., and Iglesias, A. (1996). Comparison of wheat simulation models under climate change. 2. Application of climate change scenarios, *Clim. Res.*, **7**, 271–281.

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

White, J. W., Hoogenboom, G., Kimball, B. A., and Wall, G. W. (2011). Methodologies for simulating impacts of climate change on crop production, *Field Crop Res.*, **124**, 357–368.