Uncertainties in scaling-up crop models for large-area climate change impact assessments

Frank Ewert, Lenny G. J. van Bussel, Gang Zhao, Holger Hoffmann, Thomas Gaiser, Xenia Specka, Claas Nendel, Kurt-Christian Kersebaum, Carmen Sosa, Elisabet Lewan, et al.

Chapter 10

Uncertainties in Scaling-Up Crop Models for Large-Area Climate Change Impact Assessments

Frank Ewert[1], Lenny G. J. van Bussel[1,20], Gang Zhao[1], Holger Hoffmann[1], Thomas Gaiser[1], Xenia Specka[2], Claas Nendel[2], Kurt-Christian Kersebaum[2], Carmen Sosa[3], Elisabet Lewan[3], Jagadeesh Yeluripati[4], Matthias Kuhnert[4], Fulu Tao[5], Reimund Rötter[5], Julie Constantin[6], Helene Raynal[6], Daniel Wallach[6], Edmar Teixeira[7], Balasz Grosz[8], Michaela Bach[8], Luca Doro[9], Pier Paolo Roggero[9], Zhigan Zhao[10], Enli Wang[10], Ralf Kiese[11], Edwin Haas[11], Henrik Eckersten[12], Giacomo Trombi[13], Marco Bindi[13], Christian Klein[14], Christian Biernath[14], Florian Heinlein[14], Eckart Priesack[14], Davide Cammarano[15], Senthold Asseng[15], Joshua Elliott[16], Michael Glotter[16], Bruno Basso[17], Guillermo A. Baigorria[18], Consuelo C. Romero[18], and Marco Moriondo[19]

[1]*Institute of Crop Science and Resource Conservation, Bonn, Germany*

[2]*Leibniz Centre for Agricultural Landscape Research, Müncheberg, Germany*

[3]*Swedish University of Agricultural Sciences, Uppsala, Sweden*

[4]*University of Aberdeen, Aberdeen, Scotland*

[5]*MTT Agrifood Research Finland, Mikkeli, Finland*

[6]*L'Institut national de la recherche agronomique, Auzeville, France*

[7]*Canterbury Agriculture & Science Centre, Lincoln, New Zealand*

[8]*Thünen-Institute of Climate-Smart-Agriculture, Braunschweig, Germany*

[9]*Univeristy of Sassari, Sassari, Italy*

[10]*Commonwealth Scientific and Industrial Research Organisation Land and Water, Canberra, Australia*

[11]*Karlsruhe Institute of Technology, Garmisch-Partenkirchen, Germany*

[12]*Swedish University of Agricultural Sciences, Uppsala, Sweden*

[13]*Dipartimento di Scienze delle Produzioni Agroalimentari e dell'Ambiente (DISPAA), Florence, Italy*

[14]*German Research Center for Environmental Health, Neuherberg, Germany*

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

[16]*University of Chicago, Chicago, IL, USA*

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

[18]*University of Nebraska-Lincoln, Lincoln, NE, USA*
[19]*Institute of Biometeorology — National Research Council of Italy, Florence, Italy*
[20]*Wageningen University, Wageningen, the Netherlands*

Introduction

Problems related to food security and sustainable development are complex (Ericksen *et al.*, 2009) and require consideration of biophysical, economic, political, and social factors, as well as their interactions, at the level of farms, regions, nations, and globally. While the solution to such societal problems may be largely political, there is a growing recognition of the need for science to provide sound information to decision-makers (Meinke *et al.*, 2009). Achieving this, particularly in light of largely uncertain future climate and socio-economic changes, will necessitate integrated assessment approaches and appropriate integrated assessment modeling (IAM) tools to perform them. Recent (Ewert *et al.*, 2009; van Ittersum *et al.*, 2008) and ongoing (Rosenzweig *et al.*, 2013) studies have tried to advance the integrated use of biophysical and economic models to represent better the complex interactions in agricultural systems that largely determine food supply and sustainable resource use.

Nonetheless, the challenges for model integration across disciplines are substantial and range from methodological and technical details to an often still-weak conceptual basis on which to ground model integration (Ewert *et al.*, 2009; Janssen *et al.*, 2011). New generations of integrated assessment models based on well-understood, general relationships that are applicable to different agricultural systems across the world are still to be developed. Initial efforts are underway towards this advancement (Nelson *et al.*, 2014; Rosenzweig *et al.*, 2013).

Together with economic and climate models, crop models constitute an essential model group in IAM for large-area cropping systems climate change impact assessments. However, in addition to challenges associated with model integration, inadequate representation of many crops and crop management systems, as well as a lack of data for model initialization and calibration, limit the integration of crop models with climate and economic models (Ewert *et al.*, 2014). A particular obstacle is the mismatch between the temporal and spatial scale of input/output variables required and delivered by the various models in the IAM model chain.

Crop models are typically developed, tested, and calibrated for field-scale application (Boote *et al.*, 2013; see also Part 1, Chapter 4 in this volume) and short time-series limited to one or few seasons. Although crop models are increasingly used for larger areas and longer time-periods (Bondeau *et al.*, 2007; Deryng *et al.*, 2011; Elliott *et al.*, 2014) rigorous evaluation of such applications is pending. Among the different sources of uncertainty related to climate and soil data, model parameters, and structure, the uncertainty from methods used to scale-up crop models has received little attention, though recent evaluations indicate that upscaling of crop models for climate change impact assessment and the resulting errors and uncertainties deserve attention in order to advance crop modeling for climate change

assessment (Ewert *et al.*, 2014; Rötter *et al.*, 2011). This reality is now reflected in the scientific agendas of new international research projects and programs such as the Agricultural Model Intercomparison and Improvement Project (AgMIP; Rosenzweig *et al.*, 2013) and MACSUR (MACSUR, 2014).

In this chapter, progress in evaluation of scaling methods with their related uncertainties is reviewed. Specific emphasis is on examining the results of systematic studies recently established in AgMIP and MACSUR. Main features of the respective simulation studies are presented together with preliminary results. Insights from these studies are summarized and conclusions for further work are drawn.

Methods of Scaling-Up Crop Models

Error versus uncertainty in scaling-up crop models

In crop modeling literature, the terms "error" and "uncertainty" are often used interchangeably (see Part 1, Chapter 9 in this volume). While error measures the deviation between simulated and past observations, uncertainty can be thought as a measure of the difference between simulated and likely future observations (see Part 1, Chapter 9 in this volume). While the relationship between past errors and future uncertainty is of critical importance, it is not the subject of this chapter. Here we focus on the scaling-up of crop models under past conditions that has resulted in errors that are largely unknown as many past studies have not justified, nor transparently disclosed, assumptions made in scaling models. Systematic evaluation of scaling methods should provide some understanding of the size of this error and hence help to quantify and reduce the uncertainty in applying crop models for large-area assessment in the future. We do not investigate future uncertainty in this chapter, but rather examine the errors produced, and their implications, of scaling-up crop models with historic climate/weather data. Likewise, the uncertainty of upscaling should be analyzed in conjunction with the uncertainties related to climate input data, crop model structure, and model parameters (Asseng *et al.*, 2013; Palosuo *et al.*, 2011; Rötter *et al.*, 2012), which is beyond the scope of this chapter.

Scales — resolution, extent, and coverage

The most relevant scales in crop modeling refer to space and time. Various terms are related to scales such as extent, resolution, and coverage though their exact meaning is often unclear and they are therefore defined and discussed again here following Bierkens *et al*. (2000), Ewert (2004), and van Delden *et al.* (2011; see also Fig. 1). Scale is defined as the characteristic dimension in time and space of a phenomenon or

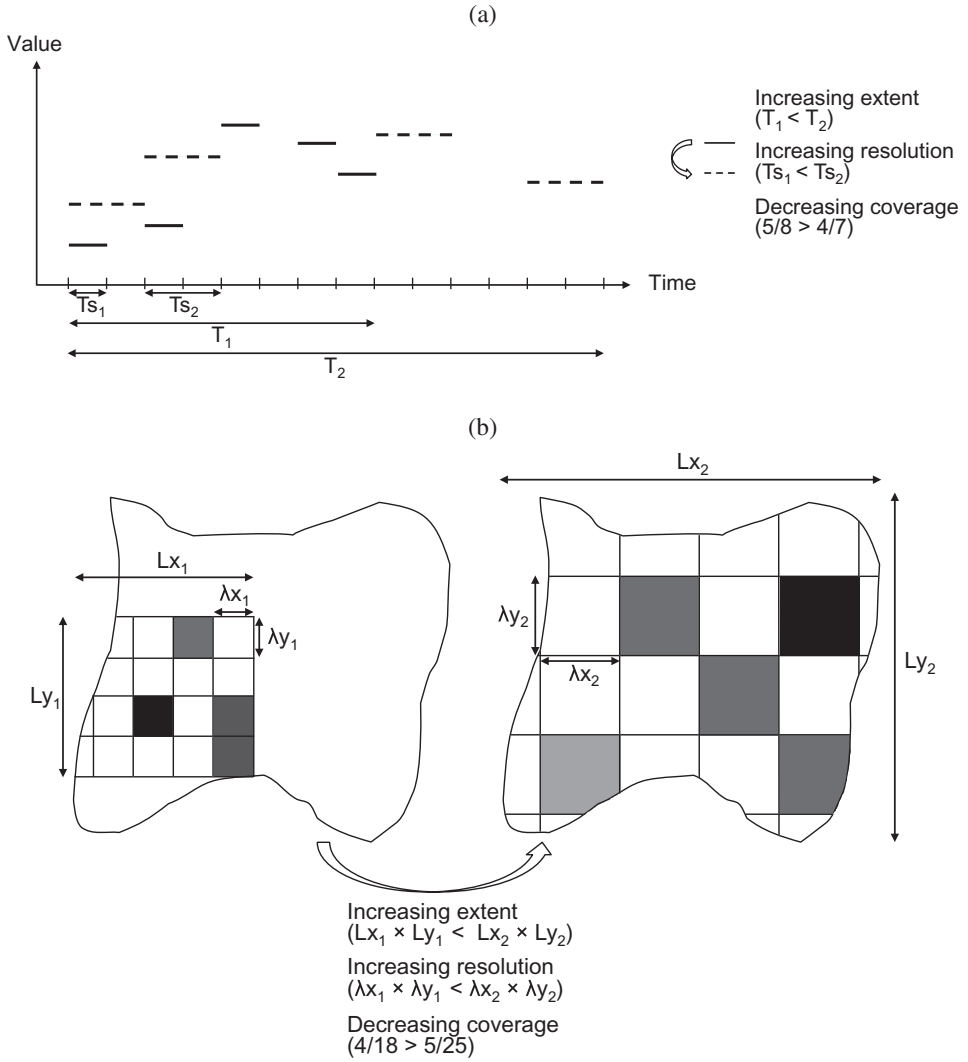

Fig. 1. Schematic representation of resolution, extent, and coverage for (a) spatial and (b) temporal scales. Increasing the spatial extent often means that the resolution becomes coarser (modified from Bierkens *et al*., 2000). Note that this study only refers to spatial upscaling.

observation, and thus dimensions and units of measurements can be assigned. In the context of this chapter, the definition of scale is colloquial (in contrast with the older cartographic definition of scale), which implies that large-scale studies deal with large extents, while small-scale studies are concerned with a small extent. Extent refers to the area and time-period over which the observations or simulations have

been made. Resolution indicates the largest spatial unit or time-period for which the variable of interest is considered to be homogeneous, i.e., only the average value of the variable is known, not the variation. A finer resolution indicates more observations per time or spatial unit, or, in the case of simulations, the use of a smaller grid size or time-step.

Coverage refers to the ratio between the sum of areas (or time-periods) for which the observations or simulations have been made and the extent (Fig. 1). Detail relates to the spatial and temporal resolution, as well as the complexity employed, in the representation of processes. Complexity is defined as the number of included relations and variables in a model. Finally, scaling is defined as transferring data and models between scales, a definition used by Blöschl and Sivapalan (1995), Ewert *et al*. (2011), and van Delden *et al*. (2011). Upscaling transfers data and models, including parameters, to a larger scale, while downscaling implies transfer of data, models, and parameters to a smaller scale. Scales should not be confused with levels of organization (Ewert *et al.*, 2011), which refer to organizational (or structural) entities such as a single plant organism, agroecosystem or farm, or the regional, national, continental, or global food system. These levels of organization have typical spatial extents and can be analyzed over specific time-periods.

Relevant processes of crop growth and yield should be modeled with sufficient functional detail, which will depend on the considered spatial and temporal resolutions. While appropriate model detail is increasingly viewed as important, the topic is still little understood (Adam *et al.*, 2011). Therefore, the use of the term "scale" implies that the application of a model refers to an area (or time-period) with a certain extent and that the model is used with a spatial and temporal resolution that should be concurrent to the resolution for which the model was originally developed. Any deviation from this will require proper evaluation (and perhaps modification of model structure, parameters, or input data) as the model is used outside the system characteristics for which it was originally developed.

Upscaling methods

While climate impact assessment studies are generally conducted on large scales, most crop models have been developed for a spatial resolution of a plot or homogenous field, with crop physiological processes typically simulated with daily temporal resolution. For some models (or processes), the resolution is higher — hours or even minutes. In theory, crop models can be applied to an infinite number of plots or homogenous fields as long as the required input data (weather, soil, and crop management) and cultivar parameters are available. The same applies to the temporal extent, i.e., the number of years for which the model is applied. In practice, large-scale model application is constrained by data availability and computing time

with relatively little consideration given to the mismatch in model process details and spatial scale of application. Therefore, methods to adjust (and likely reduce) model complexity with increasing scales are required.

Various methods are used to apply field models at larger scales, and generally are pragmatic solutions to the problem of limited data availability and/or computational power. For instance, model application in a region may be constrained to the (small) number of available weather stations from which data can be used for model simulations. Or, a model is applied in conjunction with climate change scenarios in which data are provided in grid cell format with a defined size, typically $50 \times 50\,km^2$ (e.g., Elliott *et al.*, 2014). The same applies to the temporal resolution of available climate/weather data, which may not have the resolution required for crop models (Nonhebel, 1994; van Bussel *et al.*, 2011). Such methods can be grouped according to how complexity is reduced. Accordingly, methods refer to the manipulation of (1) input data (climate, soil, management), (2) model structure, by reducing model complexity, and (3) model parameters. A comprehensive overview of scaling methods is given Ewert *et al.* (2011) and van Oijen *et al.* (2009). The most prominent methods used in climate change studies are described here.

Data aggregation, sampling, and extrapolation

Two main methods of data aggregation can be distinguished based on whether the aggregation refers to the input data or the output data after model simulation. The latter assumes that points are well sampled across a region to obtain good coverage. However, in most cases data coverage is insufficient. In the simplest case, only one so-called "representative" point is available within a region for which data (weather, soil, and crop management) are available and for which model simulations are performed. Simulation results for this point are then assumed to be representative for the entire region, i.e., simulations are extrapolated to all parts of the region (Fig. 2a). However, this is only valid if the region of concern is homogenous for all relevant input data of the model, which will rarely be the case. For heterogeneous regions, the number of sampling points should ideally be sufficient to represent the spatial heterogeneity in input data of this region. As information about the spatial variability in input data is limited, sampling points are usually randomly chosen. A more qualified way for considering heterogeneity within a region is stratified sampling (Fig. 2b). Sampling points are chosen for strata in a region that represent homogenous quantities, often climatic conditions. Hence, environmental stratifications are commonly used to support such sampling (Ewert *et al.*, 2011). However, the sampling is usually constrained by the number of available points with data. Little information is available about the relationship between change in number of sampling points and deviation in simulation results.

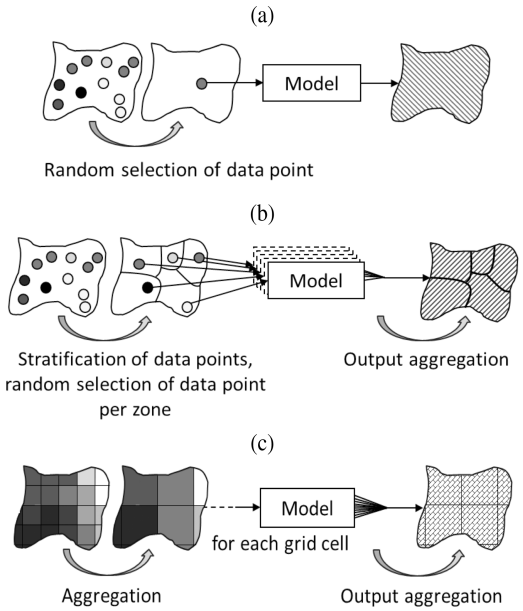

Fig. 2. Selected scaling methods often used in impact assessment studies: (a) extrapolation from a single point in the region to the entire region, (b) stratified sampling guided by an environmental stratification to sample points within relatively homogenous environmental strata, and (c) aggregation of input data to larger grid cells for which the models are run. For a more elaborated description of scaling methods see Ewert *et al*. (2011).

Alternatively, input data are aggregated and models are run with these aggregated input data (Fig. 2c), as is the case when models are used in combination with gridded climate data from climate models (Bondeau *et al.*, 2007; Deryng *et al.*, 2011; Elliott *et al.*, 2014). The main advantage of this method is that full coverage of the spatial extent of interest can be achieved. The disadvantage is that with increasingly coarse levels of aggregation, heterogeneity in input data is lost, which potentially affects simulation results as many processes in crop models are typically non-linear. Although both methods have frequently been used, uncertainties in using these methods are largely unknown. Further, possible interactions between scaling method and choice of crop model are unclear. Recent studies have pointed to the importance of using crop model ensembles in impact studies to account for the large uncertainty in crop models (Asseng *et al.*, 2013; Palosuo *et al.*, 2011; Rötter *et al.*, 2012).

Evaluating Uncertainty from Scaling Methods

Uncertainties in both methods of input and output data aggregation for model upscaling have been evaluated systematically in two comprehensive studies recently

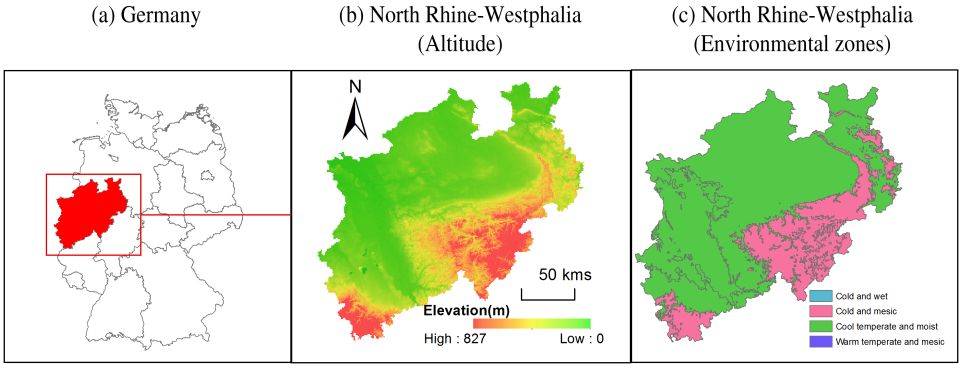

Fig. 3. Example region in (a) Germany used in scaling exercise. The region refers to the state of North Rhine-Westphalia with heterogeneous conditions of (b) altitude and (c) environmental conditions. Note that environmental zones are derived from Metzger *et al.* (2013).

launched within the international research program AgMIP (http://www.agmip.org/) and the FACCE JPI Knowledge Hub MACSUR (http://www.macsur.eu/). Simulation exercises and results are described in detail elsewhere for AgMIP (van Bussel *et al.*, 2014) and MACSUR (Hoffmann *et al.*, 2014; Zhao *et al.*, 2014). In these studies, the effect of different numbers of sampling points (AgMIP) and grid cell size (MACSUR) were investigated for winter wheat grown in the state of North Rhine-Westphalia in Germany (Fig. 3). The region was chosen because it is characterized by a considerable spatial heterogeneity in environmental (climate and soil) conditions (Fig. 3). Also, high-resolution weather and soil data were available.

Simulation exercises

Models were applied to simulate a typical winter wheat variety grown in the region. Information for model calibration only comprised a typical sowing date (October 1) and harvest date (August 1) and an indication of the regional average actual dry matter grain yield (about 7.2 t/ha) was given. Simulations were performed to obtain potential and water-limited yields. In the first step of these exercises the effect of scaling-up climate data was investigated. Other input data (soil, crop management) were kept constant. Hence, only one dominant soil type (sandy loam) with a rootable depth of 2.3 m and a total available water capacity of 429 mm was chosen for the entire region; more detailed soil information can be obtained from van Bussel *et al.* (2014). Also, crop management did not vary and was considered optimal for nitrogen fertilization and no effects of pests, diseases, or weeds were considered. Selected results presented in this chapter refer to potential growth, while results for

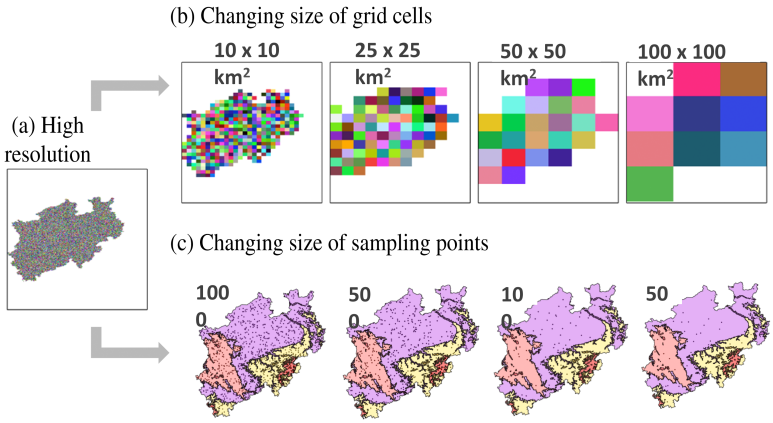

Fig. 4. Set-up of simulation experiments for evaluating deviations from (a) high-resolution simulations of (b) different grid cell sizes and (c) reduced numbers of sampling points. Details are described in van Bussel *et al.* (2014), Hoffmann *et al.* (2014), and Zhao *et al.* (2014).

water-limited conditions are shown elsewhere (Hoffmann *et al.*, 2014; van Bussel *et al.*, 2014; Zhao *et al.*, 2014).

In the AgMIP study 12 crop models were used (MONICA, APSIM, pDSSAT, HERMES, MCWLA, NWheat, SALUS, SIMPLACE<LINTUL2>, SPASS-ExpertN, STICS, CERES-wheat, and CROPGRO), a subset of those used in Asseng *et al*. (2013). Models and references are described in Asseng *et al*. (2013). Ten of the 12 models are considered in the later analysis described in this chapter. Five of these models were also used in the MACSUR study, along with six other models for a total of 11 models in this chapter (APSIM, Modified APSIM, COUP, DailyDayCent, LandscapeDNDC, EPIC, HERMES, SIMPLACE<LINTUL5>, MCWLA, MONICA, STICS). Models and references are described in Hoffmann *et al.* (2014) and Zhao *et al.* (2014).

The effect of different sample sizes (Fig. 2a, b) for climate/weather data were investigated in AgMIP; the number of sampling points per region was varied from 10, 100, 500, to 1000 points and compared with a high-resolution run for 34168 points (Fig. 4a), which refers to a 1×1 km^2 grid raster (Fig. 4c). The sampling of points was guided by the environmental stratification of Metzger *et al.* (2013). The effects of grid cell size was investigated in MACSUR by conducting simulations for the region with grid cell sizes of 10×10 km^2, 25×25 km^2, 50×50 km^2, and 100×100 km^2 (Fig. 4b). These were also compared with a high-resolution run from 100×100 km^2, i.e., 34168 cells. Observed actual yields were obtained from regional yield statistics (IT.NRW, 2014).

Impact of Scaling Method

Effect of sampling size (*AgMIP study*)

Surprisingly, differences in regional averages of simulated potential yields based on crop model ensemble means were small, depending on the number of sampling points considered (Fig. 5a). This was essentially consistent across models with differences between models for simulated yields larger than differences due to sample size (van Bussel *et al.*, 2014). The model ensemble mean also reproduced the interannual yield variability well for most years, again with small effects of sample size on temporal yield variability (Fig. 5b). As simulated yields refer to potential conditions they are about 15% to 20% higher than observed yields. Likewise, differences between models in simulating interannual yield variability were substantially larger than the small differences due to the different sample point sizes (van Bussel *et al.*, 2014). Such results suggest that with relatively few sample points (in this case ten) the relative error in regional yield simulations compared to simulations with a high density (full coverage) is small. However, our example referred to a stratified sampling case in which a sample size of ten points ensured that all environmental strata were sampled. Results may differ if points are sampled randomly without consideration of an environmental stratification (van Bussel *et al.*, 2014).

Effect of grid cell size (*MACSUR study*)

Simulation results of model ensemble means reveal some effects of changing grid cell sizes on potential yields (Fig. 5c). Average yields tend to decline as the size of the grid cells increases until a size of $50 \times 50\,km^2$ (Fig. 5c). The higher average yield at the large grid cell size of $100 \times 100\,km^2$ may be caused by the methodology of considering border grid cells, which results in a larger total area of this cell size as compared to smaller grid cell sizes. The ensemble mean of the 11 models used in this exercise also closely reproduced the interannual variability of observed regional yields (Fig. 5d). Again, model differences were considerable, as shown elsewhere (Hoffmann *et al.*, 2014; Zhao *et al.*, 2014). Grid cell size showed some effect on simulated potential yields, which changed depending on the year (Fig. 5d), but temporal patterns of ensemble means were not affected. These results suggest that for the crop, region, and models considered here, errors in simulating potential yields are comparably small for larger grid cell sizes. However, it should be noted that, also in this study, average potential yields were 15% to 20% higher than observed actual yields, which is in line with reported yield gaps for Germany.

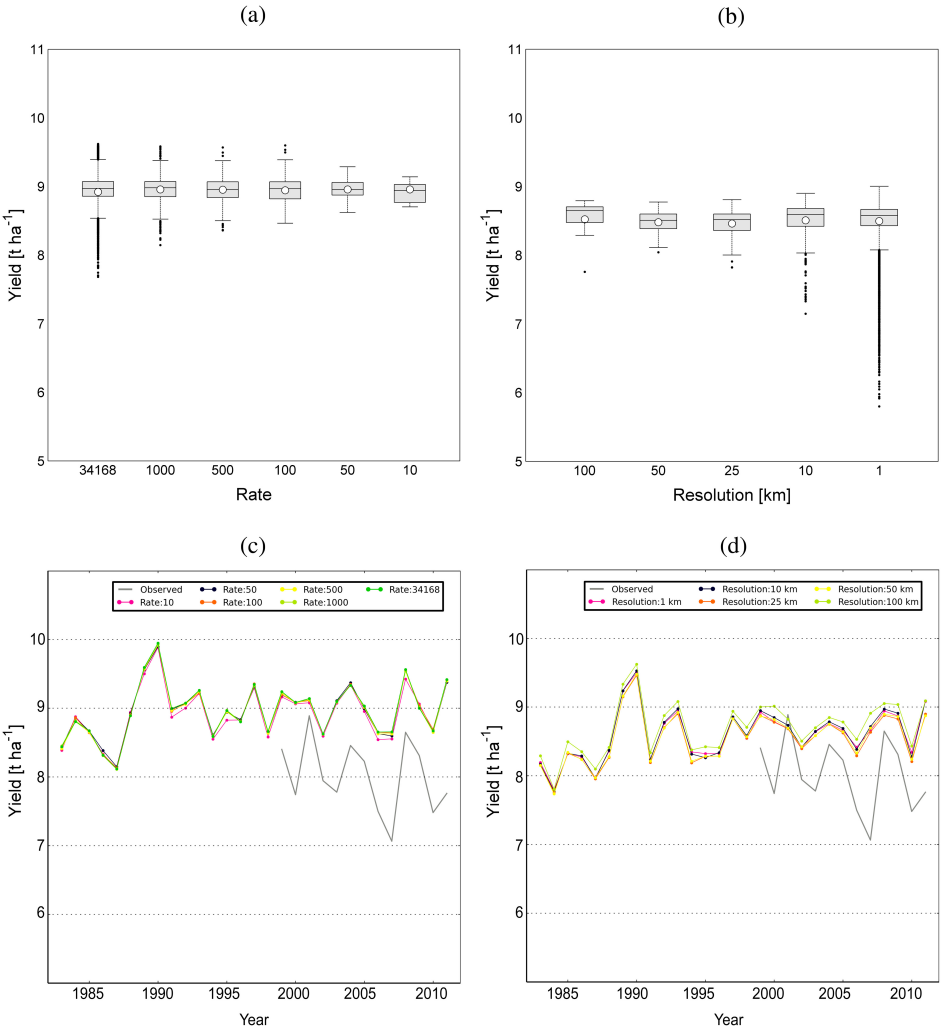

Fig. 5. Selected results from scaling exercise in AgMIP and MACSUR for simulated potential grain yield of winter wheat in North Rhine–Westphalia, Germany between 1981–2011. Results show model ensemble means of the effect of the number of sampling points (rate) on (a) 30-year yield statistics (box-and-whisker plots with mean, median, and 25th and 75th percentiles) and (b) interannual yield variability, and of the effects of grid cell size on (c) 30-year yield statistics (box-and-whisker plots with mean, median, and 25th and 75th percentiles) and (d) interannual yield variability. Gray lines in (b) and (d) refer to observed regional yield statistics. Note, different models are behind the ensemble means shown in the different panels (see the section on simulation exercises). For a more elaborated description and presentation of results see van Bussel *et al.* (2014), Hoffmann *et al.* (2014), and Zhao *et al.* (2014).

Spatial variability

Variability in simulated potential yield through space increases with increasing number of sample points (Fig. 5a) and decreasing grid cell size (Fig 5c). The effects of

climate/weather data aggregation on spatial variability have been reported earlier (Zhao *et al.*, 2015) and are also evident from this study (Zhao *et al.*, 2014). If information on the spatial variability of yields is required, studies based on few sample points or large grid cells may be insufficient. This applies particularly to areas where spatial variability in environmental conditions is high (Zhao *et al.*, 2015).

Knowledge Gaps and Future Activities

Knowledge gaps

These results (Fig. 5) are in line with a recent study (Angulo *et al.*, 2013) but contradict other studies (e.g., Hansen and Jones, 2000). The results presented here refer to potential yields and the aggregation of climate/weather data, though effects of scaling methods are more pronounced for water-limited yield (Hoffmann *et al.*, 2014; van Bussel *et al.*, 2014; Zhao *et al.*, 2014). It remains unclear how large the differences between the scaling methods with weather data become if regional variability in soil and management are considered and simulations are extended to also consider N-limited conditions. Likewise, there is no indication of to what extent results obtained in our studies can be transferred to other crops or regions. The results presented here refer to grain yield, whereas in integrated assessment studies other variables such as greenhouse gas emissions, N-leaching, N-yield will also be of interest (Ewert *et al.*, 2014). How transferable these results are across impacts variables is unknown. Particularly striking are the differences among crop models, not shown here but presented in detail elsewhere (Hoffmann *et al.*, 2014; van Bussel *et al.*, 2014; Zhao *et al.*, 2014). These which are consistent with reports from earlier studies (Asseng *et al.*, 2013) and are worth exploring in relation to the available input data for model calibration. As some interaction between scaling methods and crop models has been observed, such interactions should not be ignored in future studies. Ensemble means may be a suitable approach to avoid such interactions.

Future activities

Given these gaps, future research on better understanding the scaling-up of crop models for large-area assessments should focus on:

- Inclusion of spatial variability in soils.
- Consideration of spatial variability in management particularly for N fertilization, sowing dates, and varieties grown. A particular challenge is to understand the impact of methods to scale-up crop rotations (Teixeira *et al.*, 2014).

- Generalization of results across crops and regions.
- Adequate model structure, parameters, and use of ensemble means.

As these and other results suggest, given that uncertainty from crop models and global climate models (GCMs) is larger than from scaling methods applied to weather data input (with the qualification that soils and management variability has not been considered yet), future research should also focus on improving crop models for large-scale applications. Here the question of appropriate model detail becomes important (Adam *et al.*, 2011). To date, large-area crop models have not been developed to capture the relationships important at an aggregated regional scale. The same applies to long time-horizons.

The effect of upscaling may become more important if variability in soil and crop management is also considered. The latter is likely to be particularly relevant in regions where crop yields are more constrained by management than by climatic conditions and where spatial variability in management intensity is high. In regions with a large share of low-intensity farming systems as in Africa, crop physiological processes as considered in crop models may be less important (Webber *et al.*, 2014) than approaches of modeling effects of management. Hence, a new generation of crop models may be needed that adequately considers the range of factors and relations relevant at larger scales and that complies with the demands of integrated assessment modeling (IAM) (Ewert *et al.*, 2014).

Acknowledgements

Frank Ewert, Holger Hoffmann, and Lenny van Bussel acknowledge support from the FACCE MACSUR project (031A103B) funded through the German Federal Ministry of Education and Research. The work of Fulu Tao and Reimund Rötter was supported by the Academy of Finland, project NORFASYS (decision no 268277). The work of Carmen Sosa and Elisabet Lewan was supported by the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS) and by strategic funding for research on "Soil, Water & Landscape" from the Faculty of Natural Resources and Agricultural Sciences (Swedish University of Agricultural Sciences). Christian Biernath was financed by the Helmholtz project "REKLIM — Regional Climate Change: Causes and Effects" Topic 9: "Climate Change and Air Quality". Christian Klein was financed by the project "GeoFarm — Precision Farming Systems" Work package 10: "Soil-Plant-Modelling". Florian Heinlein was financed by the DFG Research Group FOR 1695 "Agricultural Landscapes under Global Climate Change – Processes and Feedbacks on a Regional Scale". We thank the Landesbetrieb Information und Technik Nordrhein-Westfalen

for providing regional yield data, and the German Weather Service for providing weather data.

References

Adam, M., Van Bussel, L. G. J., Leffelaar, P. A., Van Keulen, H., and Ewert, F. (2011). Effects of modelling detail on simulated potential crop yields under a wide range of climatic conditions, *Ecol. Model.*, **222**, 131–143.

Angulo, C., Rötter, R., Trnka, M., Pirttioja, N., Gaiser, T., Hlavinka, P., and Ewert, F. (2013). Characteristic "fingerprints" of crop model responses to weather input data at different spatial resolutions, *Eur. J. Agron.*, **49**, 104–114.

Asseng, S., Ewert, F., Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., Boote, K. J., Thorburn, P. J., Rötter, 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., Müller, 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., Stöckle, 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.

Bierkens, M. F. P., Finke, P. A., and de Willigen, P. (2000). *Upscaling and downscaling methods for environmental research*, Kluwer, Dordrecht.

Blöschl, G. and Sivapalan, M. (1995). Scale issues in hydrological modelling: a review, *Hydrol. Proc.*, **9**, 251–290.

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

Boote, K. J., Jones, J. W., White, J. W., Asseng, S., and Lizaso, J. I. (2013). Putting mechanisms into crop production models, *Plant Cell Environ.*, **36**, 1658–1672.

Deryng, D., Sacks, W., Barford, C., and Ramankutty, N. (2011). Simulating the effects of climate and agricultural management practices on global crop yield, *Global Biogeochem. Cycle*, **25**, GB2006.

Elliott, J., Deryng, D., Müller, C., Frieler, K., Konzmann, M., Gerten, D., Glotter, M., Flörke, M., Wada, Y., Best, N., Eisner, S., Fekete, B. M., Folberth, C., Foster, I., Gosling, S. N., Haddeland, I., Khabarov, N., Ludwig, F., Masaki, Y., Olin, S., Rosenzweig, C., Ruane, A. C., Satoh, Y., Schmid, E., Stacke, T., Tang, Q., and Wisser, D. (2014). Constraints and potentials of future irrigation water availability on agricultural production under climate change, *Proc. Natl. Acad. Sci.*, **111**, 3239–3244.

Ericksen, P. J., Ingram, J. S., and Liverman, D. M. (2009). Food security and global environmental change: emerging challenges, *Environ. Sci. Policy*, **12**, 373–377.

Ewert, F. (2004). *Modelling changes in global regionalized food production systems under changing climate and consequences for food security and environment — development of an approach for improved crop modelling within IMAGE*, Wageningen University & Netherlands Environmental Assessment Agency (MNP), National Institute for Public Health and Environment (RIVM).

Ewert, F., Rötter, R. P., Bindi, M., Webber, H., Trnka, M., Kersebaum, K. C., Olesen, J., van Ittersum, M. K., Janssen, S., Rivington, M., Semenov, M. A., Wallach, D., Porter, J. R., Stewart, D., Verhagen, J., Gaiser, T., Palosuo, T., Tao, F., Nendel, C., Roggero, P. P., Bartošová, L., and Asseng, S. (2014). Crop modelling for integrated assessment of climate change risk to food production, *Environ. Model. Software*, submitted.

Ewert, F., van Ittersum, M. K., Bezlepkina, I., Therond, O., Andersen, E., Belhouchette, H., Bockstaller, C., Brouwer, F., Heckelei, T., Janssen, S., Knapen, R., Kuiper, M., Louhichi, K., Olsson, J. A., Turpin, N., Wery, J., Wien, J. E., and Wolf, J. (2009). A methodology for enhanced flexibility of integrated assessment in agriculture, *Environ. Sci. Policy*, **12**, 546–561.

Ewert, F., van Ittersum, M. K., Heckelei, T., Therond, O., Bezlepkina, I., and Andersen, E. (2011). Scale changes and model linking methods for integrated assessment of agri-environmental systems, *Agric. Ecosyst. Environ.*, **142**, 6–17.

Hansen, J. W. and Jones, J. W. (2000). Scaling-up crop models for climate variability applications, *Agric. Syst.*, **65**, 43–72.

Hoffmann, H., Zhao, G., van Bussel, L., Enders, A., Xenia, S., Carmen, S., Jagadees, Y., Fulu, T., Julie, C., Edmar, T., Balasz, G., Luca, D., Zhigan, Z., Nendel, C., Kiese, R., Raynal, H., Eckersten, H., Haas, E., Wang, E., Kuhnert, M., Trombi, G., Bindi, M., Lewan, E., Bach, M., Kersebaum, K. C., Rötter, R., Roggero, P. P., Wallach, D., Krauss, G., Siebert, S., Gaiser, T., and Ewert, F. (2014). Effects of climate input data aggregation on modeling regional crop yields, *Clim, Res.*, in press.

IT.NRW (2014). *Statistik*. Available at: http://www.it.nrw.de/statistik/index.html. Accessed on 19 September 2014.

Janssen, S., Athanasiadis, I. N., Bezlepkina, I., Knapen, R., Li, H., Domínguez, I. P., Rizzoli, A. E., and van Ittersum, M. K. (2011). Linking models for assessing agricultural land use change, *Computer Electron. Agric.*, **76**, 148–160.

MACSUR (2014). Available at: http://www.macsur.eu/. Accessed on 21 September 2014.

Meinke, H., Howden, S. M., Struik, P. C., Nelson, R., Rodriguez, D., and Chapman, S. C. (2009). Adaptation science for agriculture and natural resource management — urgency and theoretical basis, *Curr. Opin. Environ. Sustain.*, **1**, 69–76.

Metzger, M. J., Bunce, R. G. H., Jongman, R. H. G., Sayre, R., Trabucco, A., and Zomer, R. (2013). A high-resolution bioclimate map of the world: A unifying framework for global biodiversity research and monitoring, *Global Ecol. Biogeog.*, **22**, 630–638.

Nelson, G. C., Valin, H., Sands, R. D., Havlík, P., Ahammad, H., Deryng, D., Elliott, J., Fujimori, S., Hasegawa, T., Heyhoe, E., Kyle, P., Von Lampe, M., Lotze-Campen, H., Mason d'Croz, D., van Meijl, H., van der Mensbrugghe, D., Müller, C., Popp, A., Robertson, R., Robinson, S., Schmid, E., Schmitz, C., Tabeau, A., and Willenbockel, D. (2014). Climate change effects on agriculture: Economic responses to biophysical shocks, *Proc. Natl. Acad. Sci.*, **111**, 3274–3279.

Nonhebel, S. (1994). The effects of use of average instead of daily weather data in crop growth simulation models, *Agric. Syst.*, **44**, 377–396.

Palosuo, T., Kersebaum, K. C., Angulo, C., Hlavinka, P., Moriondo, M., Olesen, J. E., Patil, R. H., Ruget, F., Rumbaur, C., Takac, J., Trnka, M., Bindi, M., Caldag, B., Ewert, F., Ferrise, R., Mirschel, W., Saylan, L., Siska, B., and Rötter, R. (2011). Simulation of winter wheat yield and its variability in different climates of Europe: A comparison of eight crop growth models, *Eur. J. Agron.*, **35**, 103–114.

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.

Rötter, R. P., Carter, T. R., Olesen, J. E., and Porter, J. R. (2011). Crop-climate models need an overhaul, *Nat. Clim. Change*, **1**, 175–177.

Rötter, R. P., Palosuo, T., Kersebaum, K. C., Angulo, C., Bindi, M., Ewert, F., Ferrise, R., Hlavinka, P., Moriondo, M., and Nendel, C. (2012). Simulation of spring barley yield in different climatic zones of Northern and Central Europe: a comparison of nine crop models, *Field Crop Res.*, **133**, 23–36.

Teixeira, E. I., Brown, H. E., Sharp, J. M., Meenken, E. D., and Ewert, F. (2014). Evaluating methods to simulate crop rotations for regional climate change assessments — a case study on the Canterbury plains of New Zealand, *Environ. Model. Software*, in press.

van Bussel, L. G. J., Ewert, F., Zhao, G., Hoffmann, H., Wallach, D., Klein, C., Biernath, C., Heinlein, F., Tao, F., Rötter, R., Cammarano, D., Asseng, S., Elliott, J., Glotter, M., Nendel, C., Basso, B., Baigorria, G. A., Romero, C. C., Kersebaum, K. C., Nendel, C., Specka, X., Gaiser, T., and Siebert, S. (2014). Spatial sampling of weather data for regional crop yield simulations, *Global Ecol. Biogeog.*, in press.

van Bussel, L. G. J., Müller, C., van Keulen, H., Ewert, F., and Leffelaar, P. A. (2011). The effect of temporal aggregation of weather input data on crop growth models' results, *Agric. Forest Meteorol.*, **151**, 607–619.

van Delden, H., van Vliet, J., Rutledge, D. T., and Kirkby, M. J. (2011). Comparison of scale and scaling issues in integrated land-use models for policy support, *Agric. Ecosyst. Environ.*, **142**, 18–28.

van Ittersum, M. K., Ewert, F., Heckelei, T., Wery, J., Alkan Olsson, J., Andersen, E., Bezlepkina, I., Brouwer, F., Donatelli, M., Flichman, G., Olsson, L., Rizzoli, A. E., van der Wal, T., Wien, J. E., and Wolf, J. (2008). Integrated assessment of agricultural systems - A component-based framework for the European Union (SEAMLESS), *Agric. Syst.*, **96**, 150–165.

van Oijen, M., Thomson, A., and Ewert, F. (2009). *Spatial upscaling of process-based vegetation models: An overview of common methods and a case-study for the U.K.*, StatGIS GeoInformatics for Environmental Surveillance., Milos Island, Greece, 17–19 June 2009., p. 6.

Webber, H., Gaiser, T., and Ewert, F. (2014). What role can crop models play in supporting climate change adaptation decisions to enhance food security in Sub-Saharan Africa?, *Agric. Syst.*, **127**, 161–177.

Zhao, G., Hoffmann, H., van Bussel, L., Enders, A., Xenia, S., Carmen, S., Jagadees, Y., Fulu, T., Julie, C., Edmar, T., Balasz, G., Luca, D., Zhigan, Z., Nendel, C., Kiese, R., Raynal, H., Eckersten, H., Haas, E., Wang, E., Kuhnert, M., Trombi, G., Bindi, M., Lewan, E., Bach, M., Kersebaum, K. C., Rötter, R., Roggero, P. P., Wallach, D., Krauss, G., Siebert, S., Gaiser, T., and Ewert, F. (2014). Impact of weather data spatial resolution on simulation of cropping systems: a comparison of models, production systems and crops, *Clim. Res.*, in press.

Zhao, G., Siebert, S., Rezaei, E. E., Yan, C., and Ewert, F. (2015). Demand for multi-scale weather data for regional crop modelling, *Agric. Forest Meteorol.*, 156–171.