INSTITUT FÜR NUTZPFLANZENWISSENSCHAFTEN UND RESSOURCENSCHUTZ
LEHRSTUHL FÜR ALLGEMEINEN PFLANZENBAU

Re-Solution - understanding the uncertainty in regional applications of crop models

Inaugural-Dissertation

zur

Erlangung des Grades

Doktor der Agrarwissenschaften
(Dr. Agr.)

der

Landwirtschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelm-Universität Bonn

von

Carlos Lenín Angulo Villacís

aus

Quito - Ecuador

Referent: Prof. Dr. Frank Ewert

Korreferent: Prof. Dr. Reimund Rötter

Korreferent: Dr. Thomas Gaiser

Tag der mündlichen Prüfung: 19.12.2014

Erscheinungsjahr: 2015

ABSTRACT

Crop models, typically developed for field-scale applications, are increasingly used to regionally assess the impacts of climate change and adaptation on agricultural production. This is due to their ability to consider dynamic interactions between genotype, environment, and management factors and variabilities herein, as particularly relevant for climate. Uncertainties emerging from the scale change and data constraints when analyzing and further utilizing the results of regional crop model applications are largely unclear. A thorough analysis distinguishing between the three main uncertainty sources attributable to regional crop model application: (a) model structure, (b) parameters and (c) input data is crucial to develop robust assessment approaches and modelling tools that can support policy decisions concerning adaptation of agricultural systems to climate variability and change. The present thesis offers a systematic analysis, particularly on the two last points mentioned above. On the one hand, we investigated the calibration quality and strategy which concerns parameters (b) and input data (c). On the other hand, we assessed the influence of resolution of input data (c) on regional modelling results for crop models differing in model structure and detail (a). Three specific studies were designed to increase understanding on these issues:

1. In a continental simulation study (EU25) the influence of considering the sub-regional differences in environmental and management conditions on the parameter estimation process of a model were investigated. Three different calibration strategies were tested: (i) calculation of phenology parameters only, (ii) consideration of both phenology calibration and a yield correction factor and (iii) calibration of phenology and selected growth processes. The third strategy, i.e. taking into consideration sub-regional differences of model parameters related to crop growth in addition to crop phenology resulted in the best agreement between simulated and observed yield at the European scale. However, since accurate calibration of crop growth and development parameters requires data which are presently scarce in the required quality and resolution for entire Europe, the use of a yield correction factor after phenology calibration (strategy 2) might be still meaningful and is advised as the preferred strategy.

2. A regional study in Jokioinen, representing an important barley producing region in South-West Finland, was undertaken in order to systematically analyse the influence of aggregation of weather data on yield simulations. The responses of four crop models of different complexity to five weather data aggregation levels (Weather station, 10 km x 10 km, 20 km x 20 km, 50 km x 50 km , and 100 km x 100 km) were compared. Differences between models were larger than the effect of the chosen spatial weather data resolution. Models showed different characteristic ‘fingerprints’ of simulated yield frequency distributions independent of the resolution used for yield simulation. Additionally, using one model (SIMPLACE<LINTUL-SLIM>) the effect of aggregation of model input *versus* model output data was assessed. Results showed that aggregating weather data had a smaller effect on the yield distribution than aggregating simulated yields which caused a

deformation of the model fingerprint. For the studied region and period, changes in the spatial resolution of weather input data introduced less uncertainty to the simulations than the use of different crop models. However, it was concluded that more evaluation will be required for other regions with a higher spatial heterogeneity in weather conditions. Also, it would be necessary to undertake a similar study considering input data related to soil and crop management.

3. Thus, a complementary study to the weather data aggregation was formulated for soil input data and undertaken in the State of North-Rhine Westphalia in Germany. This comprised a systematic analysis of the influence of three different spatial soil data resolutions on simulated regional yields and simulated total growing season evapotranspiration. The resolutions used corresponded to soil maps of the scales: 1 : 50 000; 1 : 300 000 and 1 : 1 000 000. The responses of four crop models of different complexity were compared. In contrast to the weather data resolution study, a model using the Richards approach (DAISY) was considered. This was to particularly pay due attention to different modelling approaches with respect to the simulation of soil water dynamics. Differences between models were again larger than the effect of the chosen spatial soil data resolution. Three main causes were identified as possible explanations for the low influence of soil data resolution on yield simulations: a) the high precipitation amount in the region b) the methods applied to calculate water retention properties and c) the method of data aggregation. No characteristic "fingerprint" between sites, years and resolutions could be found for any of the models.

After an integrated analysis and synthesis of the results of all three studies in this thesis, one main conclusion was that data collection and data administration protocols should be implemented at regional and larger scale (e.g. within projects such as MACSUR and AgMIP). Additionally, the utilization of various crop models differing in complexity and approaches of modelling relevant processes should become common practice for large area impact assessment studies since the uncertainties introduced by the model choice have been shown in this study to be more important than the uncertainties caused by the input data resolution. Nevertheless, since the areas chosen for studies 2 and 3 had a large influence on the results, for regions with more spatial variation in weather than in study 2, and with more variable and erratic rainfall than in study 3, results might look differently and therefore both a model ensemble approach and proper scaling methods might be needed. As well, the crop modelling community is appealed to make an effort to develop accurate and consistent model parameter estimation methodologies and strategies.

ZUSAMMENFASSUNG

Pflanzenwachstumsmodelle, die typischerweise für Ertragseinschätzungen kleiner Flächen Verwendung finden, werden zunehmend auch genutzt, um den Einfluss von Klimawandel auf landwirtschaftliche Produktion sowie deren Anpassung an verschlechterte Anbaubedingungen auf regionaler Ebene zu untersuchen. Grund dafür ist, dass diese Modelle dynamische Wechselwirkungen zwischen Genotyp, Umwelt und darin enthaltenen Management-Faktoren sowie die Variabilität dieser drei Einflussgrößen berücksichtigen können.

Dabei ist weitgehend unklar, welche Unsicherheiten aus der veränderten Skalierung und der eingeschränkten Datenlage entstehen, wenn man die Ergebnisse regionaler Anwendungen von Pflanzenwachstumsmodelle analysiert und weiter verwendet. Um robuste Untersuchungsmethoden und Modellierungswerkzeuge zu entwickeln, die als Grundlage für politische Entscheidungen über die Anpassung landwirtschaftlicher Systeme an Klimavariabilität und -wandel dienen können, ist eine sorgfältige Analyse unabdingbar, die zwischen den drei Hauptunsicherheitsquellen regionaler Anwendungen von Pflanzenwachstumsmodellen unterscheidet: (a) Modellstruktur, (b) Parameter und (c) Eingangsdaten. Die vorliegende Arbeit bietet eine systematische Analyse insbesondere der beiden letztgenannten Punkte. Sie untersucht zum einen die Kalibrierungsgüte und -strategie, was Parameter (b) und Eingangsdaten (c) betrifft. Zum anderen ermittelt sie den Einfluss der Eingangsdatenauflösung auf die Ergebnisse regionaler Modellierungen für Pflanzenwachstumsmodelle, die sich in ihrer Struktur und Genauigkeit (a) unterscheiden. Drei spezifische Studien wurden entworfen, um in diesen Punkten zu einem besseren Verständnis zu gelangen.

1. In einer kontinentweiten Simulationsstudie (EU25) wurde untersucht, inwiefern die Miteinbeziehung der subregionalen Unterschiede in Umwelt- und Managementbedingungen den Parameterschätzungsprozess eines Modells beeinflusst. Drei verschiedene Kalibrierungsstrategien wurden erprobt: (i) alleinige Berechnung von Phänologieparametern, (ii) Berücksichtigung der Phänologiekalibrierung und eines Ertragskorrekturfaktors und (iii) Kalibrierung von Phänologie und ausgewählten Wachstumsprozessen. Die dritte Strategie, d.h. die Berücksichtigung von sub-regionalen Unterschiede pflanzenwachstumsbezogener Modellparameter sowie von Pflanzenphänologie, ergab die größte Übereinstimmung zwischen simuliertem und tatsächlichem Ertrag auf europäischer Ebene. Da jedoch für eine exakte Kalibrierung von Pflanzenwachstum und Entwicklungsparametern Daten nötig wären, die zurzeit kaum in der benötigten Qualität und Auflösung für ganz Europa vorhanden sind, könnte es dennoch von Bedeutung sein, nach der Phänologiekalibrierung (Strategie 2) einen Ertragskorrekturfaktor zu verwenden.

2. Eine regionale Studie in Jokioinen, einem wichtigen Gersteanbaugebiet in Südwestfinnland, wurde durchgeführt, um den Einfluss der Aggregierung von Wetterdaten auf Ertragssimulationen

systematisch zu analysieren. Dazu wurde das Verhalten von vier Pflanzenwachstumsmodellen unterschiedlicher Komplexität auf fünf verschiedenen Wetterdatenaggregierungsstufen (Wetterstation, 10 km x 10 km, 20 km x 20 km, 50 km x 50 km und 100 km x 100 km) verglichen. Die Unterschiede zwischen den Modellen waren größer als die Auswirkungen der jeweiligen Wetterdatenauflösung. Jedes Modell zeigte einen charakteristischen „Fingerabdruck“ in Bezug auf die Wahrscheinlichkeitsverteilung der simulierten Erträge, unabhängig von der für die Ertragssimulation verwendeten Auflösung. Zusätzlich wurde für ein Modell (SIMPLACE<LINTUL-SLIM>) der Effekt der Aggregierung von Input- im Vergleich zu derjenigen von Outputdaten untersucht. Die Ergebnisse zeigten, dass die Aggregierung von Wetterdaten eine geringere Auswirkung auf die Ertragsverteilung hatte als die Aggregierung simulierter Erträge, die eine Verformung des Fingerabdrucks der Modelle zufolge hatte. Für die untersuchte Region brachten Veränderungen in der räumlichen Auflösung der Wetterdaten im Untersuchungszeitraum weniger Unsicherheit in die Simulationen ein als der Gebrauch unterschiedlicher Pflanzenwachstumsmodelle. Es wurde jedoch festgestellt, dass weitere Evaluationen für andere Regionen mit einer größeren räumlichen Heterogenität der Wetterbedingungen vonnöten sein werden. Außerdem bestünde der Bedarf, eine ähnliche Studie zu den Inputdaten in Bezug auf Boden- und Pflanzenmanagement zu durchzuführen.

3. Daher wurde eine ergänzende Studie zur Wetterdatenaggregierung für Bodeninputdaten formuliert und in Nordrhein-Westfalen durchgeführt. Dies umfasste eine systematische Analyse des Einflusses dreier verschiedener räumlicher Bodendatenauflösungen auf simulierte regionale Erträge und simulierte Evapotranspiration der Gesamtwachstumsperiode. Die verwendeten Bodendatenauflösungen entsprachen Bodenübersichtskarten der Maßstäbe: 1 : 50 000; 1 : 300 000 und 1 : 1 000 000. Die simulierten Erträge von vier verschieden Wachstumsmodellen wurden verglichen. Anders als bei der Studie, die den Einfluss von Wetterdatenauflösung untersuchte, wurde hier zusätzlich ein Modell, das den Richards-Ansatz verwendet (DAISY) verwendet. Somit wurden unterschiedliche Modellansätze im Bezug auf Bodenwasserhaushaltberechnungen berücksichtigt. In den Ergebnissen waren die Unterschiede zwischen den Modellen erneut größer als die Auswirkungen der jeweiligen Bodendatenauflösungen. Drei Hauptgründe wurden als mögliche Erklärungen dafür vorgeschlagen: a) die hohe Niederschlagsmenge im untersuchten Gebiet, b) die zur Berechnung der hydraulischen Bodeneigenschaften angewandten Methoden und c) die verwendete Aggregierungsmethode. Charakterisierende Fingerabdrücke für Subregionen, Jahren oder Auflösungen konnten für kein Modell gefunden werden.

Nach einer integrierenden und synthetisierenden Analyse der drei genannten Studien in dieser Dissertation ist eine daraus entstehende Hauptschlussfolgerung: Es müssen Protokolle zum Sammeln und Verwalten von Daten auf regionaler und großflächiger Ebene eingeführt werden (wie zB. in Projekten wie MACSUR und AgMIP). Da die Ergebnisse dieser Dissertation ergaben, dass die Unsicherheiten, die durch die Modellwahl entstehen, größer sind als die Unsicherheiten, die durch die Auflösung von Inputdaten verursacht werden, sollte der Einsatz von mehreren Modellen, die sich in Bezug auf Komplexität und Art der Modellierungsansätze verschiedener Prozesse unterscheiden, ein unverzichtbarer Bestandteil der großflächigen Auswirkungseinschätzungen von Klima(wandel) auf Erträge werden. Dennoch sind die Ergebnisse von Studien 2 und 3 von den Eigenschaften der gewählten Regionen abhängig. Daher könnten für Regionen mit höher räumlicher Variabilität der Wettereigenschaften als in Studie 2 und mit heterogenerer und unregelmäßigerer Niederschlagsverteilung als in Studie 3 die jeweiligen Ergebnisse unterschiedlich ausfallen. In diesem Falle empfiehlt sich der Einsatz von Modellensembles und passenden Skalierungsmethoden. Zugleich sollen die Modellierungsforschergruppen gemeinsam an der Entwicklung von genaueren und konsequenteren Parametereinschätzungsmethoden und -strategien arbeiten.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

General Introduction

1. General introduction

1.1 Crop models: from field to region

1.1.1 Yield assessment tools

Mechanistic crop growth models (further on referred to as crop models) are relatively simple mathematical representations of a crop and its physiological processes, used to study crop growth and development (Penning de Vries et al., 1989) and typically have been developed for small spatial extents, i.e. plots or fields. Some research efforts have aimed to develop crop models to be explicitly applied at regional level (e.g. Bondeau et al., 2007; Challinor et al., 2004; Tubiello and Fischer, 2007). The regional characteristics of these models are not always evident and therefore they are not considered in the present work. Instead, it is assumed and evaluated in this thesis that field scale crop models might be useful as regional assessment tools if the uncertainties caused by scale change can be reported and quantified.

Statistical approaches have also been applied to assess agricultural productivity at different spatial extents (e.g. Lobell et al., 2008). However, the utilization of crop models is regarded as more advantageous when process-based explanations of the behaviour of crop systems are required. Also, statistical models are confronted with the problem of confounding (e.g. Bakker et al., 2005) The ability of crop models to consider dynamic interactions between genotype, environment, and management factors makes them a powerful tool and they are increasingly used to regionally assess the impacts of climate change and adaptation on agricultural production (e.g. Challinor et al., 2010; Therond et al., 2011; White et al., 2011). The regional applicability of crop models is a promising field of research (e.g. Adam et al., 2012; De Wit et al., 2010; Hansen and Jones, 2000; Moen et al., 1994; Reidsma et al., 2009a; Therond et al., 2011). There are especially two areas that require critical reflection in this respect.

First, the validity domain of the majority of crop models is limited to the spatial extent at which they have been developed and validated, i.e. plot or field scale (e.g. Boogaard et al., 1998; Jones et al., 2003a; Spitters, 1990; Van Ittersum et al., 2003a; Williams et al., 1983). Thus, when applying crop models at larger spatial extents, emerging scale change issues have to be handled properly in order to produce useful results from regional assessments of crop productivity (Ewert et al., 2011b).

Second, and closely related to the first area, since the original application of the majority of crop models is limited to small spatial extents, they require detailed information about the crop's environment (weather and soil) and management as input data (Faivre et al., 2004). Due to the high spatial and temporal variability of the mentioned variables and the relative scarcity of observations with adequate spatial density, challenges in manipulating model input data to apply model for larger spatial extents need to be addressed. In this regard, it is essential to develop coherent strategies that allow and facilitate the utilization of available input data for the regional application of crop models (Ewert et al., 2011b; Hansen and Jones, 2000; Launay and Guerif, 2005; Leenhardt et al., 2006).

In the present work, the terms *scale, resolution and aggregation* will be consistently utilized following the definitions given by O'Neil and King (1998), Faivre *et al*. (2004), and Ewert *et al.* (2006). Accordingly, **scale** is used as a synonym of spatial extent and refers to the spatial dimension of a phenomenon studied. The term **resolution** refers to the ratio between the area covered by observations and the total area considered by a study (extent). Finally, **aggregation** refers to the sum, count or average of the information at a (lower) biological/biophysical organization level to reach a higher hierarchical level. The same terminology can be applied to the temporal dimension which however is not considered as it is not the subject of this thesis.

1.1.2 State of the art of regional applications

During the early 1990s, first global assessments of climate change impacts on food security were performed utilizing plot and field scale crop models as tools for assessing climate change impacts on rice, wheat, soybean and maize at global scale (e.g. Parry et al., 1999; Rosenzweig and Parry, 1994). In these studies, experimental data from over 100 sites were used to simulate the possible crop responses to global warming and raising atmospheric CO_2 concentration. The simulation results based on experimental stations data were extrapolated for important production regions all over the world and used in conjunction with economical models to build scenarios of the possible effects of global climate change in global food production (Rötter et al., 2013b). In the following years an increasing number of research groups began to utilize crop models as regional assessment tools (e.g. Rötter and Van Diepen, 1994), especially (but not only) in the context of climate change (White et al., 2011). Consequently, it became necessary to establish simulation protocols (e.g. Moen et al., 1994; Rosenthal et al., 1998) and to suggest approaches to deal with the scale change issues and input data limitations inherent to the regional application of crop models (e.g. Faivre et al., 2004; Hansen and Jones, 2000). As well, model inter-comparison exercises, at the small scale, were undertaken in order to identify systematic errors and further improve crop models (e.g. Ewert et al., 2002; Jamieson et al., 1998; Porter et al., 1993).

During the last two decades some work has been done in order to identify and estimate the uncertainties emerging from the regional application of crop models (e.g. Easterling et al., 2001; Easterling et al., 1998; Mearns et al., 2001; Mearns et al., 2003; Niu et al., 2009; Olesen et al., 2007; Rötter et al., 2013b; Rötter et al., 2011b; Trnka et al., 2007). Nevertheless, only few examples in the literature can be found, of studies investigating explicitly and systematically the uncertainties in regional crop model applications. In these studies, the influence of the temporal and spatial resolution of weather and soil input data on crop phenology and yields received special attention (Folberth et al., 2012; Nendel et al., 2013; Olesen et al., 2000; Van Bussel et al., 2011a; Van Bussel et al., 2011b; Wassenaar et al., 1999). Studies investigating the uncertainties in regional crop model applications introduced by crop model parameters (Therond et al., 2011) and by the model equations (Adam et al., 2012) are almost an exception. Recent studies have proven the usefulness of utilizing several crop models (multi-model ensembles) as a mean to evaluate and reduce uncertainty in regional crop modelling applications (Asseng et al., 2013; Palosuo et al., 2011; Rötter et al., 2012b).

1.2 The sources of uncertainty in regional crop modelling applications

Uncertainty is defined by Walker et al. (2003) as "any departure from the unachievable ideal of complete determinism". Following the line of Walker et al. (2003), the term uncertainty in the present work refers to the accumulated uncertainty reflected in the model outputs caused by propagation and accumulation of the uncertainties in the model structure, model inputs and model parameters.

The crop modelling community has gained awareness of the lacking attention on the uncertainty when using crop models regionally, especially in the context of climate change assessment. As a result, two international initiatives have addressed the challenge of assessing and reporting uncertainties in climate change impact projections on agriculture and food security, and of improving crop models in order to reduce some of the uncertainty. These are: the Agricultural Model Inter-comparison and Improvement Project (AgMIP, Rosenzweig et al., 2013) and the European MACSUR (www.macsur.eu), the first Knowledge Hub launched by the Joint Research Progamming Initiative on Agriculture, Food Security and Climate Change (FACCE-JPI) (www.faccejpi.com).

Clearly, uncertainty reporting and quantification is crucial to develop robust assessment approaches and modelling tools that can support policy decisions concerning food security and adaptation of agricultural systems to climate change at different scales (Rötter et al., 2011a). Thus, it becomes necessary to consider the uncertainty related to the two critical areas mentioned above (section 1.1.1) when analysing and further utilizing the results of regional crop model applications. Uncertainty analysis should also distinguish between the three main sources: model structure, parameters and input data.

1.2.1 Model structure

Crop models consist of a series of equations which represent the soil-plant-atmosphere system (Faivre et al., 2004). In the common case, these equations have been developed for field scale applications. Consequently, uncertainty emerges when applying crop models regionally since it is not clear if the described processes might be appropriate for larger spatial extents and additionally, due to the scale change, new processes might become even more important than the ones taken into consideration by the original model's equations (Ewert et al., 2006). On the one hand, it is argued that the level of detail in which processes are described by field scale crop models might be too demanding considering that data at larger scales are usually more aggregated and less detailed (Ewert et al., 2011b). On the other hand, it has been shown that oversimplification of the important processes such as light utilization might lead to the omission of important relationships (Adam et al., 2011). In general, to avoid unnecessary uncertainty introduction, it is suggested to utilize equations which consider mechanisms immediately related to the yield-determining processes (Challinor et al., 2009a), ideally at the respective scale.

1.2.2 Model parameters

The mathematical equations, which are integral part of crop models, contain coefficients which are commonly known as parameters (Faivre et al., 2004). The process of parameter estimation, also called calibration, plays a decisive role on the quality of model results (Wallach et al., 2010). Apart from the error introduction which is inherent to the parameter estimation process of any model (Palosuo et al., 2011), the regional application of crop models might act as an additional source of uncertainty .

Since the processes described by the majority of crop models are highly detailed, it is difficult to find measured data for larger spatial extents which are qualitatively and quantitatively sufficient to be used for parameter estimation (Ewert et al., 2006). An alternative to overcome the data scarcity might be generating a set of regionalized parameters adapted to the spatial extent at which the model is applied (Ewert et al., 2011b).

Until present, the parameter values for regional crop model applications are not estimated but usually obtained from the literature, assuming that they can be uniformly applied over large regions (e.g. De Wit et al., 2010; Harrison and Butterfield, 1996). Nevertheless, literature values are often outdated and do not consider the possible yield improvements in new crop varieties, even when derived from comprehensive analysis of experimental studies (Rötter et al., 2011a). In this respect, it has been recommended to re-estimate the parameter values for regional applications in order to improve the capability of the models to capture spatial yield variability between sub-regions (Reidsma et al., 2009a). In the context of regional agricultural productivity assessments, especially in relation to the impacts of climate change and variability, the accurate depiction of yield variability between regions plays an important role (Challinor et al., 2009a; Hansen and Jones, 2000; Reidsma et al., 2009a). Few studies have suggested a methodology of parameter estimation which considers these differences. A first step in this direction has been taken by Therond *et al.* (2011). The authors propose to use region-specific factors for 12 European regions in order to correct simulations of phenological stage for regional differences. They conclude that this strategy substantially improves yield simulations in comparison to using the same phenological parameter set for all regions. They also strongly recommend proceeding the same way with growth parameters for obtaining more accurate simulation of growth in biomass and yield.

1.2.3 Model input data

As mentioned above, the application of crop models at larger spatial extents is severely hindered by their detailed data requirements. Several methods have been suggested to overcome this problem (e.g. Faivre et al., 2004; Hansen and Jones, 2000; Leenhardt et al., 2006). They are mainly based on some form of data aggregation (e.g. Mearns et al., 2003; Van Der Velde et al., 2009). For a comprehensive description of such methods to manipulate models and data for larger scale crop model applications the reader is referred to Ewert *et al.* (2011b). There is an increasing interest in explicitly assessing the uncertainties inherent to the process of data aggregation and manipulation (e.g. Niu et al., 2009; Olesen et al., 2000; Van Bussel et al., 2011a; Van Bussel et al., 2011b). An important question regarding this field of research is to which extent the choice of spatial resolution of environmental input data (soil, weather) influences model simulations (e.g. De Wit et al., 2005; Folberth et al., 2012; Mearns et al., 2001; Nendel et al., 2013). However, the effect of spatial aggregation of input data on simulated yields has only been partially assessed until now and demands more systematic analyses. Also, possible interactions between model related uncertainties (Asseng et al., 2013; Palosuo et al., 2011; Rötter et al., 2013b; Rötter et al., 2012b) and uncertainty due to scaling are in need of attention.

1.3 General objective and research questions

In response to the need of more insight into the critical aspects inherent to the application of crop models at larger spatial extents, the overall objective of the present PhD thesis is to systematically address the uncertainties emerging from the regional application of crop models. For this purpose I aim to answer the three following research questions.

Question 1 (Q1).- *What is the relevance of considering region-specific differences in the calibration of a crop model at large scale?*

This question focuses on examining different calibration strategies for simulating spatial and temporal yield variability.

Question 2 (Q2).- *What are the effects of changes in the spatial resolution of weather input data on the simulation results of diverse crop models?*

Question 3 (Q3).- *What are the effects of changes in the spatial resolution of soil input data on the simulation results of diverse crop models?*

The changes in the spatial resolution mentioned in Q2 and Q3 concern different data aggregation levels based on a high spatial resolution data set which is systematically manipulated to obtain lower resolution input data sets. Since the influence of the method of aggregation is not the focus of this question, simple averaging in the case of weather input data and sampling of the most representative units in the case of soil input data are used.

The effects of the aggregation levels of input data on the simulation results are evaluated in terms of the influence of data resolution on simulation statistics including frequency distributions of model outputs.

The crop models taken into consideration use different approaches to simulate crop growth and development, and describe those physiological processes at different degrees of complexity.

1.4 Study setting

In order to answer Q1 a continental yield simulation study considering 25 member countries of the European Union (EU 25) was established. This choice was made based on the availability of two extensive databases: the European SEAMLESS database (Andersen et al., 2010; Van Ittersum et al., 2008) which provides soil and weather information and the JRC/MARS Crop Knowledge Base (JRC, 1998) containing yearly sowing and harvest data for grain maize, potato, sugar beet, winter barley and winter wheat.

The yield simulation study dealing with Q2 was undertaken in the Yläneenjoki region, a rather small (approx. 4000 km^2) and well established barley producing region in south-western Finland. The reason to consider this area was the access to a climate data set from the Finish Meteorological Institute available for whole Finland with a 10 km x 10 km grid cells resolution (Venäläinen et al., 2005), which served as highest data resolution used as basis for the stepwise aggregation of weather input data. An special advantage of the selected area was the availability of crop data for 400 to 600 parcels from the MYTAS database (Palva et al., 2001). Based on this information a supplementary analysis of the influence of aggregation on the distributions of observed yield data was possible.

Finally, a yield simulation study involving seven counties in the Federal State of North-Rhine Westphalia was set up in order to answer Q3. The access to the exceptionally detailed soil data information, a soil map at a scale of 1:50 000 (BK50, 2004) provided by the Geological Service of North-Rhine Westphalia, justifies choosing this region. The soil map contains information about the distribution of approximately 7000 soil units for the entire State which are used as starting point for the stepwise aggregation of soil input data.

1.5 Structure of the thesis

The thesis comprises 5 chapters. This first chapter is the general introduction, while chapters 2 to 4 deal with the three research questions in the order mentioned above. **Chapter 2** compares three different model calibration strategies using: i) region-specific phenology parameters but only one parameter set for the EU 25, ii) both phenology calibration and a region-specific final yield correction factor and iii) calibration of phenology and region-specific parameters for selected growth processes. The effects of the tested calibration strategies are assessed and the best performing strategy is used to estimate the impacts of climate change combined with increasing CO_2 concentration and technology development on yields. In **Chapter 3** the effect of spatial aggregation of weather data on regional yields simulated by four crop models is investigated. The frequency distributions of yields simulated using five weather data resolutions are compared and the differences between them caused by the choice of the resolution for each model and between models are evaluated. Additionally, the effects of model input *vs.* model output aggregation on simulated yield distributions are investigated as well as the effects of aggregating observed yields as compared to the aggregation of simulated yields. Similarly, **Chapter 4** explores the importance of aggregation of soil input data when simulating regional yields using four crop models. Three spatial resolutions are tested and the frequency distributions of simulated yields and simulated total growing season evapotranspiration are compared to assess the influence of soil input data resolution on model results. Moreover, the behaviour of the four crop models with respect to soil input data resolution is analysed to gain insights into the uncertainty of model simulations. Finally, in **Chapter 5,** the main findings of the PhD thesis are summarized and discussed. Emerging research questions related to the regional application of crop models are proposed and a number of recommendations are given concerning the assessment of uncertainty in regional climate change impact projections for agriculture.

Chapter 2

Implication of crop model calibration strategies for assessing regional impacts of climate change in Europe

This chapter has been published as:

Angulo, C., Rötter, R., Lock, R., Enders, A., Fronzek, S. and Ewert, F., 2013. Implication of crop model calibration strategies for assessing regional impacts of climate change in Europe. Agric. For. Meteorol., 170: 32-46.doi:10.1016/j.agrformet.2012.11.017

2. Implication of crop model calibration strategies for assessing regional impacts of climate change in Europe

2.1 Introduction

Despite the persisting challenge in scaling up detailed information on crop growth and development from the field to the regional and higher level (Ewert et al., 2011b; Faivre et al., 2004; Hansen and Jones, 2000), process-based crop simulation models (hereafter referred to as crop models) are a commonly used tool for large area impact assessment of climate variability and change on crop yields (e.g. Challinor et al., 2010; Olesen and Bindi, 2002; Parry et al., 2004; Rötter et al., 2011a; Xiong et al., 2008). In contrast to statistical models (e.g. Lobell et al., 2008) crop models provide process-based explanations of systems behaviour to changes in the environment. Furthermore crop models are able to consider dynamic interactions between environment, genotype and management factors, which justifies their application in projecting impacts of climate change and adaption on agricultural crop production. However, large area applications of these models are often hindered by limited data availability for model calibration and testing and extensive computing time. Most large scale applications of crop models have some way of considering the spatial variability of input data such as climate, soil characteristics and management practices, often through some form of data aggregation (Fischer et al., 2005; Mearns et al., 2001; Van Der Velde et al., 2009; Wassenaar et al., 1999). However, only few attempts have been made to quantify errors related to the method of input data aggregation (Olesen and Bindi, 2002; Van Bussel et al., 2011a; Van Bussel et al., 2011b). Even less information is available about the importance of model calibration for large area applications with aggregated and scarcely available data from observations.

It is well recognised that model calibration is indispensable to improve the accuracy of yield estimations in climate change studies (Jagtap and Jones, 2002; Wolf et al., 1996) and that it has implications for the overall reliability of the model simulations (Challinor et al., 2009a). With a few exceptions (e.g. Challinor et al., 2004) crop parameters are usually not subjected to calibration, but they are obtained from the literature assuming that they can be uniformly applied over large regions. For example, De Wit *et al.* (2010) used a multiple crop parameter set based on field experiments in the Netherlands, UK and Belgium to simulate crop yields in Europe with the WOFOST model. Harrison and Butterfield (1996) considered variety-specific phenology parameters for major wheat growing regions in Europe, but did not calibrate any other growth parameters. In fact, recent model applications have considered regional differences in phenological development, but only few examples are known were the regional variability of growth parameters is considered. For instance, Xiong *et al.* (2008) used a cross-calibration procedure that explicitly accounted for variety characteristics and proposed variety-specific parameter sets for each of 16 major rice producing zones in China in order to capture the spatial variability of regional yields. In global studies, models such as GAEZ (Fischer et al., 2005) and LPGmL (Bondeau et al., 2007) consider region-specific parameters for phenology and some selected growth processes, however, it remains a challenge to develop parameters with a common approach that are comparable and have been extensively tested for all regions.

Recently, Therond *et al.* (2011) argued that crop model parameters can only be derived from field experiments where growth and development processes have been measured. Furthermore, they suggest that aggregated data from regional statistics are insufficient to derive parameters for crop models as these have been originally developed for field-level applications based on field experiments. Instead, the authors propose an approach to calibrate the phenology module of the crop model APES (Donatelli et al., 2010; Van Ittersum et al., 2008) for 12 European regions using correction factors which were calculated based on the differences between simulated and observed phenology dates. When these region-specific factors where applied to correct simulated dates of phenological stages, simulations of grain yield improved substantially in comparison to yield simulations with only one parameter set for all regions. However, Therond *et al.* (2011) also stressed that calibration of phenology is not sufficient to reproduce observed yields across regions in Europe.

Some efforts are known where parameters have been derived from comprehensive analysis of experimental studies, which, however, date back decades ago (Boons-Prins et al., 1993). In a recent analysis of European-wide simulations with the WOFOST model (Boogaard et al., 1998) using these parameters, Reidsma *et al.* (2009a) concluded that a re-calibration of crop-growth-related parameters could improve the model's capability in capturing spatial (regional) yield variability. Model parameters apparently referred to old varieties and crop improvement was not considered, a phenomenon common to many of the widely used crop models (Rötter et al., 2011a). This also applies to impact assessment studies were crop and management improvement over time is hardly accounted for. Very few examples are known (e.g. Ewert et al., 2005; Hermans et al., 2010) where estimations of climate change effects on crops are combined with scenario dependent assumptions about changes in agro-technology development affecting yield potential and the yield gap (Lobell et al., 2009).

Spatially explicit and comprehensive studies assessing the influence of climate change on agricultural yields in Europe using crop models are scarce (Harrison and Butterfield, 1996; Trnka et al., 2011). The few attempts that have been made, e.g. Van Der Velde *et al.* (2009) for pan-European rapeseed production, have highlighted the importance of considering the regional differences in climate change assessments. Acknowledging the large data, parameter, and output uncertainties when using process-based models in regional yield prediction, Ewert *et al.* (2005) suggested a simple empirical approach to estimate crop productivity under climate change in Europe that accounts for regional yield variability and temporal changes due to crop and management improvement. However, effects of climatic variability could not be assessed with this approach.

None of the above-mentioned studies has investigated the possible effects of model calibration for crop phenology, growth and yield-related parameters on large area simulations and possible implications for estimations of climate change impacts. Accordingly, in this study we aim to investigate the importance of a region-specific model calibration for simulations of five crops across 25 member countries of the European Union (EU25). We do not question the need for a region-specific calibration of phenology parameters as this is already commonly accepted. The focus of our study is on the calibration of growth processes and yield. We assume that region-specific parameters for growth and yield will improve yield simulations, including their spatial and temporal variability. More specifically, we compare three different ways of calibrating a crop model (further referred to as calibration strategies) using (i) calculated region-specific phenology parameters and one growth-influencing parameter set for all regions in Europe, (ii) consideration of both phenology calibration and a region-specific final yield correction factors, and (iii) calibration of phenology and region-specific parameters for three selected growth processes. Simulations are performed with LINTUL2 (Van Ittersum et al., 2003b) combined with a calibration algorithm implemented in the modelling interface LINTUL-FAST for 533 climate zones (Andersen et al., 2010) across EU25. We also test to which extent the different calibration strategies affect estimations of climate change impacts on crop yields. Finally, we consider the best performing calibration strategy to estimate the impacts of climate change in combination with increasing atmospheric CO_2 concentration and technology development on yields. These impact projections are performed stepwise to understand the individual contributions of climate change, increasing atmospheric CO_2 concentration and technology development on estimated yield changes.

2.2 Materials and Methods

2.2.1 Model Description

Our modelling activities are based on the crop model LINTUL2 for potential and water-limited conditions (Farré et al., 2000; Spitters and Schapendonk, 1990; Van Ittersum et al., 2003b). As the original model LINTUL2 simulates phenology only for spring crops it was extended with a phenology model as used in APES together with LINTUL2 (Adam et al., 2012). LINTUL2 was combined with a search algorithm, (see 2.5) to calibrate parameters. LINTUL2 equipped with the mentioned calibration algorithm was developed in the modelling interface FAST which allows fast simulations for large numbers of spatial units and years for which temporal model performance becomes a critical issue. The resulting model combination LINTUL-FAST is used in this study. LINTUL2 considers effects of climate including limited water supply as described in (Farré et al., 2000; Spitters and Schapendonk, 1990). It has been used in numerous climate change studies (e.g.

Ewert et al., 1999; Hijmans, 2003; Van Oijen and Ewert, 1999; Wolf and Oijen, 2002). Different from other model versions (Ewert et al., 1999; Rodriguez et al., 2001; Van Oijen and Ewert, 1999; Wolf and Oijen, 2002) for the present study a simple representation of the effects of increased atmospheric CO_2 concentration (denoted as $[CO_2]$) on biomass production was considered using the relationship between $[CO_2]$ and radiation use efficiency proposed by Stockle *et al.* (1992):

$$RUE_e = (100)([CO_2])/[-[CO_2] + b_l \exp(-b_2 [CO_2])] \quad (1)$$

where RUE_e is Radiation use efficiency in g MJ^{-1} and $[CO_2]$ represents the atmospheric $[CO_2]$ in ppm. The values assigned to the parameters b_l and b_2 are 6928 and 0.0014 respectively, and correspond to a moderate increase of RUE due to atmospheric $[CO_2]$ elevation from 350 to 600 ppm (Stockle et al., 1992). This relationship was assumed for all crops except for grain maize which is a C4 plant and presents no (≤1%) stimulation of photosynthesis at elevated (≥600 ppm) atmospheric $[CO_2]$ (Leakey et al., 2009). The second effect of $[CO_2]$ on biomass production is to reduce crop transpiration. A linear diminution of transpiration up to 10% for all crops was taken into consideration when the atmospheric $[CO_2]$ reaches 700 ppm (Ewert et al., 2002; Kruijt et al., 2008). In addition, for calibration strategy 3, where one specific calibrated value of RUE is utilized for each climate zone and which is used for the performed climate change impact simulations, the fertilization effect of elevated atmospheric $[CO_2]$ was calculated using a correction factor:

$$RUE_n = RUE_{en} * RUE_{c0} / RUE_{e0} \quad (2)$$

Where RUE_n is the corrected RUE value for any future year n depending on CO_2 concentration, RUE_{en} is the RUE value obtained for the correspondent n year when applying equation (1), RUE_{c0} is the RUE value obtained from calibration strategy 3 for ambient CO_2 concentration and RUE_{e0} is the RUE value under ambient CO_2 concentration when applying equation (1).

2.2.2 Weather data

Weather data were obtained from the SEAMLESS database (Janssen et al., 2009; Van Ittersum et al., 2008) for 533 climate zones in EU25 (Andersen et al., 2010; Janssen et al., 2009) for the period 1983-2006. A climate zone is defined a spatial unit that combines NUTS-2 (Nomenclature of Territorial Units for Statistics) regions and Environmental Zones (EnZ) (Metzger et al., 2005). Data included daily rainfall (mm d^{-1}), maximum air temperature (°C), minimum air temperature (°C), global solar radiation (MJ m^{-2} d^{-1}), wind speed (m s^{-1}) and vapour pressure (hPa). Evapotranspiration (mm d^{-1}), was available from the observed database where it was calculated with the Penman-Monteith formula as applied by Allen *et al.* (1998a).

2.2.3 Soil data

Soil characteristics at the level of AgriEnvironmental Zones (AEnZ) (Hazeu et al., 2010), a further refinement of the climatic zones were also available from the Pan European SEAMLESS database (Andersen et al., 2010; Van Ittersum et al., 2008). Six different soil types were defined according to topsoil organic carbon levels (Hazeu et al., 2010). However, in this study only the dominant soil type per AEnZ, i.e. the soil type covering the largest area in each AEnZ, was considered and aggregated to the level of NUTS-2 administrative regions for which yield statistics were also available.

2.2.4 Crop data

2.2.4.1 Crop phenology

Yearly sowing and harvest dates for grain maize, potatoes, sugar beet, winter barley and winter wheat were obtained from the JRC/MARS Crop Knowledge Base for 233 NUTS-2 regions across Europe (JRC, 1998). However, due to missing values in some NUT2 regions and years, these dates were averaged to the level of EnZ across Europe. Subsequently, the obtained sowing and harvest dates for the 13 EnZs were disaggregated again to the climate zones. These data of sowing and harvest dates were then used for the calibration of LINTUL-FAST.

2.2.4.2 Crop yields

Annual yields were available for NUTS-2 regions from 1983 to 2006 from the EUROSTAT database (Eurostat, 2010). For Germany, data gaps were noticed and filled with data from the Federal Office of Statistics of Germany (Destatis, 2010). Other data gaps could not be filled, so that consistent data for the entire period were not available for all regions (Section 2.3.2, Figure 4d).The yield data were the basis for the calibration exercise of LINTUL-FAST.

2.2.5 Model calibration

2.2.5.1 Calibration criteria

LINTUL-FAST uses an optimization brute-force search algorithm for the calibration of crop phenology, three biomass production parameters and the yield correction factor. The targeted parameters were determined by the minimum root mean square error RMSE between simulated and observed data given by:

$$RMSE(\theta_s - \theta_o) = \sqrt{[\sum^{n}_{i=0}(x_{s,i} - x_{o,i})^2/n)]} \qquad (3)$$

where *s* is simulated and *o* is observed yield, θ is a yield data vector and *x* is a yield data point. The calibration algorithm was set up to search for the best value for each considered parameter (i.e. minimising RMSE) within a maximum of eight iterations. Tests have shown that larger numbers of search iterations improve parameter values only marginally.

2.2.5.2 Calibration procedure and strategies

In this study we test the effect of three different strategies of calibrating the crop model LINTUL-FAST,

(1) Region-specific parameters of phenological development only,

(2) Region-specific phenology parameters and a correction factor for yield estimations.

(3) Region-specific phenology parameters and calibration of selected growth parameters instead of a yield correction factor.

Strategy 1: calibration for phenological development

For all three calibration strategies temperature sums for 533 climate zones of EU25 were calculated using aggregated observed crop phenology data for the stages sowing and maturity, and the historical weather data at climate zones level. Values of temperature sums were calculated from sowing to anthesis and from anthesis to maturity for each climate zone based on available data (see Section 2.2.4). Due to the uncertainty regarding the variation of base temperature among genotypes and development stages (McMaster et al., 2008), one base temperature value was considered for each crop and applied for all climate zones (Yin and Van Laar, 2005). Growth influencing parameters were not calibrated in this strategy, thus one set of growth parameters for each crop was used for all regions across Europe. Table 1 presents an overview of the main crop growth influencing parameter values considered for each crop.

Table 1. Default parameters as used in simulations of calibration strategies 1 and 2. (RUE=radiation use efficiency, SLA=specific leaf area, DT=drought tolerance).

	Parameter			References
Crops	RUE ($g\ Mj^{-1}$)	SLA (m^2g^{-1})	DT (-)	
Winter wheat	2.8	0.028	0.3	Garcia et al., 1988; Yin and Van Laar, 2005*
Winter barley	2.9	0.031	0.3	Goyne and Hare, 1993, Adam pers. com+
Potato	2.7	0.033	0.4	Spitters and Schapendonk, 1990
Sugar beet	3.5	0.02	0.4	Jaggard et al., 2003
Maize	3.8	0.022	0.2	Farré et al., 2000

* Xinyou and Laar, 2005 provide values for SLA for all crops.

+ Adam provided values for DT for each crop

Strategy 2: calculation of yield correction factors

This strategy extends from strategy 1, region-specific phenology parameters are used (strategy 1) in addition to a yield correction factor. This yield correction factor was calculated for each climate zone based on minimising RMSE between observed and simulated yields from 1983 until 2006, as follows:

$$\lambda = \sum s_i o_i / \sum s_i^2 \qquad (4)$$

where λ is the yield correction factor, *s* is the simulated and *o* the observed yield in an *i*-zone. The obtained yield correction factor for each climate zone was applied, to all years in this climate zone. The available yields statistics were de-trended to exclude yield increases resulting from technology development. For this purpose, yield trends were calculated for each climate zone by fitting a linear regression line through the correspondent observed yields, as proposed by Ewert et al. (2005). The yearly yields in each climate zone were then de-trended by adding or subtracting the correspondent value of the slope of the linear regression. The yield trends were explicitly considered in the scenario analysis (see section 2.2.6). No calibration of growth parameters was performed and one set of growth parameters was used for all regions in Europe.

Strategy 3: calibration of growth parameters

Selected growth parameters were calibrated using observed crop yields from 1983-2006 which were de-trended as described above under strategy 2. The calibration referred to three parameters, (i) radiation use efficiency (RUE), (ii) specific leaf area (SLA) and (iii) drought tolerance (DT). It was assumed that these parameters represent main variety differences in leaf area index and thus light capturing, light conversion to biomass and drought sensitivity. The incorporated calibration algorithm in LINTUL-FAST allows for a simultaneous search of the 3 parameters:

$$\|Sim(RUE_i^n, SLA_j^n, DT_k^n)-o\| \leq \|Sim(RUE_i, SLA_j, DT_k)-o\| \text{ for all } i,j,k \qquad (5)$$

where,

$RUE_i=RUE_0-RUE_0*r+(RUE_0*r*2)/7*(i-1)$

$SLA_j=SLA_0-SLA_0*r+(SLA_0*r*2)/7*(j-1)$

$DT_k=DT_0-DT_0*r+(DT_0*r*2)/7*(k-1)$

i^n, j^n,k^n = selected values

i,j,k=1,2...8

r=allowed percentage of variation of the parameter.

The total variation of each parameter was limited to 45% (r=0.45) of its default value. Three successive iterations were undertaken to search for the best parameter values since preliminary tests showed that more than three iterations did not further minimize RMSE significantly. For the first search iteration RUE_0, SLA_0 and DT_0 were set to the default values of each parameter (Table 1). For the second interaction RUE_0, SLA_0 and DT_0 were replaced with the resulting values of RUE_i, SLA_j and DT_k from the first iteration. For the third iteration, the values of RUE_0, SLA_0 and DT_0 were replaced with the resulting values of RUE_i, SLA_j and DT_k from the second iteration. No yield correction factor was considered in this strategy. Thus, instead of a yield correction factor, one set of growth parameters was provided for each climate zone and was applied to all years for which data were available in this climate zone.

2.2.6 Scenario analysis

2.2.6.1 Climate change scenarios

The scenario analysis considered changes in climate, atmospheric [CO_2] and technological development and compared a baseline scenario (1983-2006) with future scenarios for the period 2041-2064.

Data from an ensemble of simulations with 15 coupled atmosphere-ocean General Circulation Models (GCMs) for three emission scenarios (10 GMCs with SRES B1 forcing, 15 with A1B and 14 with A2, Nakicenovic *et al.*, 2000) were downloaded from the CMIP3 archive (Meehl et al., 2007) for those variables required for crop modelling. A subset of the following seven scenarios was selected to span the range of changes in temperature and precipitation by the mid-21st century:

- SRES A1B 15-model ensemble mean (15GCM A1B) – this provides a central projection of the changes with respect to all variables.
- Pattern-scaled SRES B2 15-model ensemble mean (15GCM B2) – all changes of the A1B ensemble mean are reduced by a scaling factor obtained from a simple climate model to emulate difference in the forcing.
- BCCR_BCM2_0/SRES B1 (BCCR B1) – less warming consistent across all regions and seasons.
- MIROC3.2(hires)/SRES A1B (MIROC A1B) – more warming consistent across all regions and seasons.
- CCCMA-CGCM3.1/SRES A2 (CGCM A2) – wet in northern Europe.
- MIROC3.2(hires)/SRES B1 (MIROC B1) – wet in central Europe.
- GISS_MODEL_E_H/SRES A1B (GISS A1B) – dry in central and northern Europe.

Simulated monthly changes between the periods 1980-1999 and 2040-2059 were calculated from the GCMs for all required variables, averaged for the 533 regions of the observed weather data, interpolated to daily deltas and added to the observed time series. In this simple delta-change approach, possible changes in inter-annual or daily variability were not considered. All scenario data were checked to provide physically plausible values, for further details see Ewert *et al.* (2011a).

The present and future atmospheric [CO_2] were based on emission scenarios taken from the Special Report on Emissions Scenarios (SRES -Nakicenovic and Swart, 2000): A1B, B1, A2 and B2. According to these projections average [CO_2] concentrations for the period 2041-2064 are 531.6, 534, 486.6 and 478.3 ppm for the SRES scenarios A1B, A2, B1 and B2, respectively.

2.2.6.2 Technology development

The importance of considering technology development in climate change impact assessments studies has been stressed by several authors (Challinor et al., 2009a; Ewert et al., 2005; Rötter et al., 2011a; Semenov and Halford, 2009). Here we use the approach described in Ewert *et al.* (2005) to estimate yield changes due to improved varieties and crop management. In this approach, historic yield trends are used as a basis to extrapolate yields into the future. The extrapolated trends are, however, modified depending on scenario specific assumptions about breeding progress to increase potential yields and crop management to reduce the yield gap (Ewert et al., 2005). In this study we used the same technology parameters to correct the historic yield trends as described in Ewert *et al.* (2005). Importantly, historic trends were calculated for the period 1983-2006 for each NUTS-2 region and disaggregated to the climate zone. Thus, all climate zones in one NUTS-2 region use the same historic yield trend. Calculated scenario-specific yield changes due to technology development were then used to correct simulated yields under climate change and increased [CO_2].

2.2.7 Simulation

The calibrated model LINTUL-FAST was used to simulate five annual crops, i.e. winter wheat, winter barley, potato, sugar beet and grain maize for Europe (EU25) for the baseline period 1983 - 2006. Future crop yields were simulated for the 24 year period centred around 2050 (2041-2064) for the 7 climate change scenarios described above (Section 2.2.6.1). In order to analyse separately the effects of climate, increased atmospheric [CO_2] and technology development, each scenario was run in three steps. First, simulations considered the influence of climate change on yields only. The next step included also the effect of increased [CO_2]. Finally, in the third step, the influence of technology development was considered in addition to the effects of climate change and increased [CO_2].

2.3 Results

2.3.1 Effect of calibration strategies on simulations

2.3.1.1 Baseline conditions

Table 2. Summary statistics of observed and simulated yields of 5 crops using three calibration strategies applied in Europe over 24 years (1983 to 2006).

Crop	Statistic	Yield			
		Observed	Strategy 1	Strategy 2	Strategy 3
Winter wheat	Mean	3.84	5.14	3.56	3.75
	St. Dev	1.78	2.45	2.01	1.83
	RMSE		2.36	1.10	0.70
	CV(RMSE)		0.56	0.26	0.17
	R^2		0.26	0.72	0.86
Winter barley	Mean	3.37	6.87	3.27	3.31
	St. Dev	1.34	2.02	1.40	1.33
	RMSE		3.94	0.72	0.59
	CV(RMSE)		1.17	0.21	0.16
	R^2		0.42	0.75	0.81
Potato	Mean	5.11	5.49	4.67	4.93
	St. Dev	1.99	2.83	2.42	2.11
	RMSE		2.65	1.54	1.13
	CV(RMSE)		0.52	0.30	0.21
	R^2		0.12	0.63	0.73
Sugar beet	Mean	11.21	6.16	8.99	10.13
	St. Dev	3.19	3.68	6.45	4.27
	RMSE		6.50	5.97	3.45
	CV(RMSE)		0.58	0.53	0.31
	R^2		0.06	0.25	0.42
Grain maize	Mean	5.97	8.41	5.53	5.54
	St. Dev	2.03	3.16	2.32	2.31
	RMSE		4.33	1.94	1.81
	CV(RMSE)		0.73	0.32	0.30
	R^2		-0.06	0.37	0.44

Abbreviations: St. Dev=Standard deviation, RMSE=root mean square error, CV(RMSE)=coefficient of variation of RMSE [RMSE/mean], CV=coefficient of variation, R^2=coefficient of determination, Strategy 1= calculation of phenology parameters only, Strategy 2= consideration of both phenology calibration and a yield correction factor, Strategy 3=calibration of phenology and selected growth processes.

An overview of the effects of the three applied calibration strategies is presented in Table 2 summarized for all crops and regions over 24 years (from 1983 to 2006).

Simulations considering regional differences in phenology only (i.e. strategy 1, section 2.2.5.2) resulted in large differences between simulated and observed yields for all crops as depicted exemplarily for winter wheat in Figure 1a. With the exception of winter barley, no relationship between simulated and observed yields could be obtained. Considering region-specific correction factors (in addition to phenology parameterisation) for yield simulations (i.e. strategy 2, section 2.2.5.2) noticeably improved simulation results for all crops but to a different extent depending on the crop (Table 2). However, there was still disagreement between observed and simulated yields which could be further reduced by applying a more extended calibration of growth parameters (i.e. strategy 3, section 2.2.5.2) as evident for winter wheat (Figure 1c), but also for other crops (Table 2).

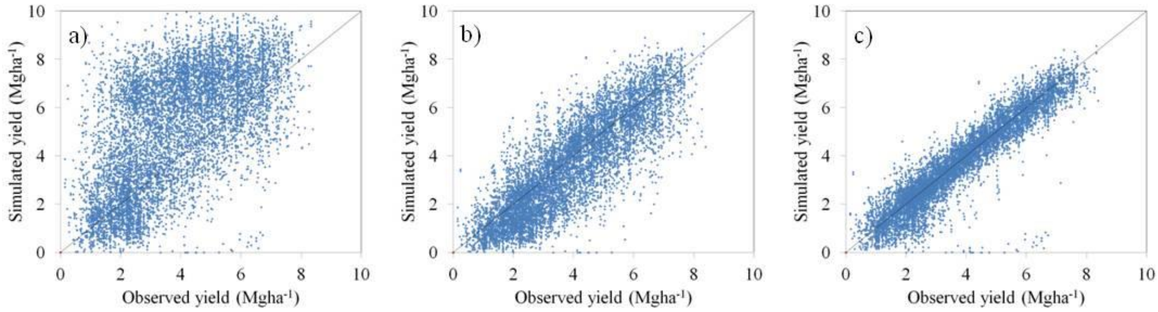

Figure 1. Comparison between observed and simulated yields from three calibration strategies, (a) phenology only, (b) using a yield correction factor, and (c) an extended calibration of selected growth parameters of winter wheat for 533 climate zones in Europe in the period from 1983 to 2006. See text for explanation of calibration strategies.

2.3.1.2 Climate change effects

Further analysis revealed that the simulated climate change effects depend on the calibration strategy used (Figure 2 and Figure 3). For instance, comparison of calibration strategies for wheat with respect to the simulated yield difference between the climate change scenarios (here: 15 GCM A1B) and the baseline showed different relationships depending on which strategies were compared (Figure 2).

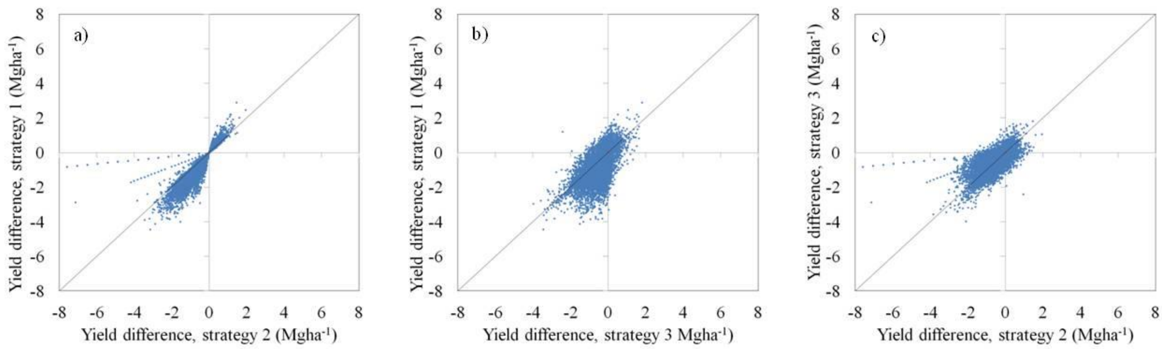

Figure 2. Relationships between calibration strategies for absolute yield differences between simulations from baseline and climate change scenario 15 GCM A1B for (a) basic calibration (strategy 1) vs. yield correction factor (strategy 2), (b) basic calibration (strategies 1) vs. extended calibration (strategy 3) and (c) extended calibration (strategy 3) vs. yield correction factor (strategy 2), for winter wheat in 533 climate zones in Europe over 24 years (1983-2006). See text for explanation of calibration strategies. (The fact that most points are located in the bottom left quadrant points out that all calibration strategies predict on average a negative effect of climate change).

For most crops except for sugar beet application of strategies 2 and 3 resulted in smaller simulated yield differences between a climate change scenario and the base line as compared to strategy 1 (Figure 3).

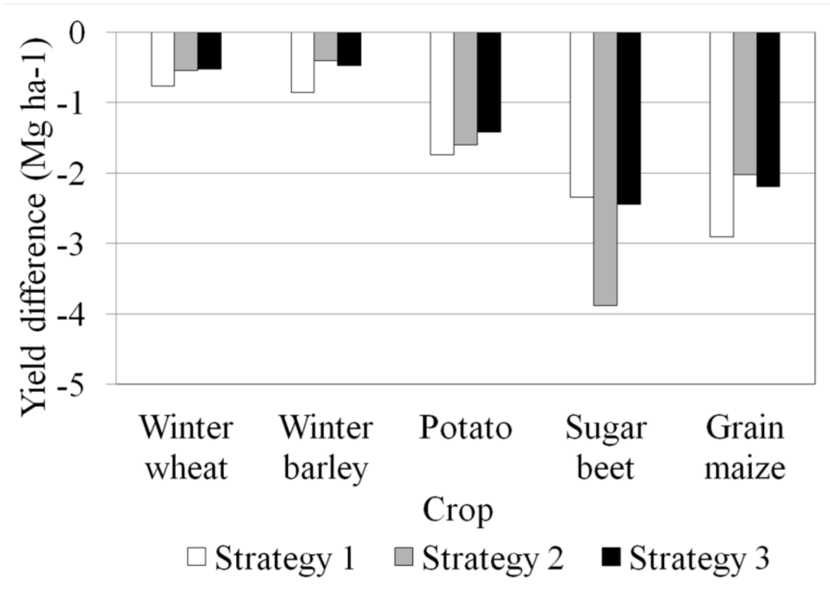

Figure 3. Absolute yield differences between simulations from baseline and climate change scenario 15 GCM A1B for three calibration strategies and five crops. Data represent averages over 533 climate zones in Europe and 24 years (1983-2006). See text for explanation of calibration strategies.

2.3.2 Spatial and temporal variability

As evident from Figure 1 and Table 2 model calibration considering growth parameters (strategy 3) provided the best agreement between observed and calibrated yields. Thus, further analysis was restricted to this calibration strategy.

A comparison of the simulated spatial pattern of wheat yields averaged over the 24 year period (1983-2006) with the observations over the same time period showed the expected good agreement between simulated and observed data (Figure 4a, b). Observed high productivity in regions of Central and Western Europe (France, Belgium, The Netherlands and Germany) and low productivities in regions of the Mediterranean countries (Spain, Italy and Greece) were also simulated by the model (Figure 4a, b).

This good agreement between simulated and observed data is not surprising as spatial differences are considered in the calibration through region-specific parameters. However, there were differences in the simulation results for individual regions. RMSEs were particularly high in regions of southern and parts of northern Europe (Figure 4c). Further analysis revealed that relative RMSEs were high in regions where observed yields were low (Figure 4e). This may point to a limitation of the present strategy to calibrate on actual yields, particularly when the gap between actual and potential yield is large due to drought and factors not accounted for in the model such as pests, diseases and weeds. It should also be noted that the number of years available for calibration differed among regions (Figure 4). However, this did not explain regional differences in model accuracy (Figure 4). These results suggest that temporal variability of yields within each region was better reproduced in regions where observed yields were high. At the aggregated EU25 scale simulated yields agreed well with the observed temporal yield variability (see Figure 5). However, there was some effect due to the incomplete time series of yield data in several regions. If only data from regions were considered for which yield data were available from 22-24 years (i.e. about 40% of all regions) the model accuracy slightly declined but showed still some agreement with the observed temporal variability (Figure 5b).

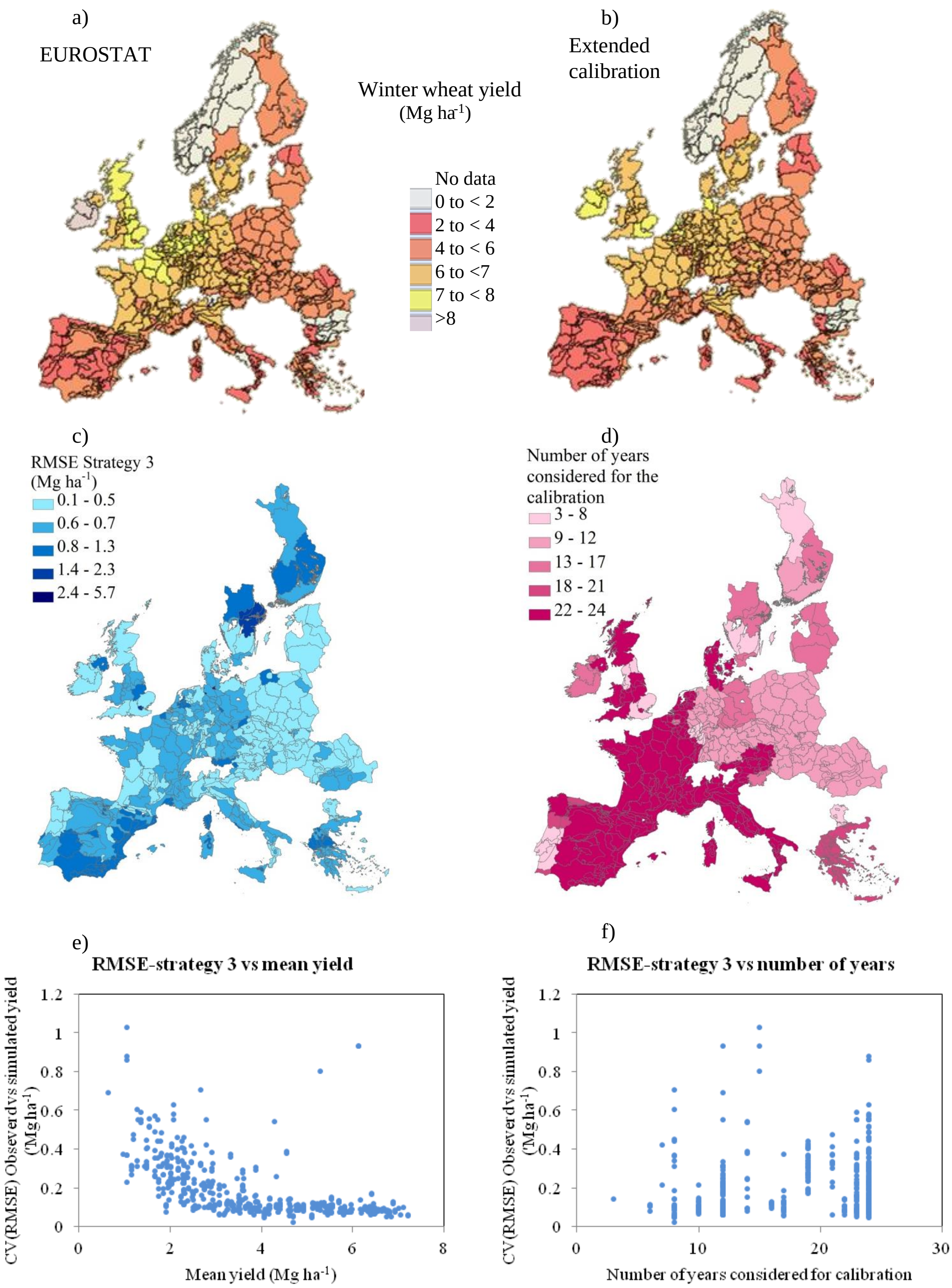

Figure 4. Spatial patterns of model calibration results for strategy 3 for winter wheat for 533 climate zones in Europe between 1983 and 2006, considering (a) observed and (b) simulated winter wheat yields (Mg ha^{-1}), (c) regional RMSE between observed and simulated winter wheat yields and (d) number of years per region with observed winter wheat yields used for model calibration. Relationships between e) CV(RMSE) of observed and simulated yield and observed mean yield, and between f) CV(RMSE) of observed and simulated yield and the number of years considered for model calibration. CV(RMSE) is the coefficient of variation of RMSE [RMSE/mean].

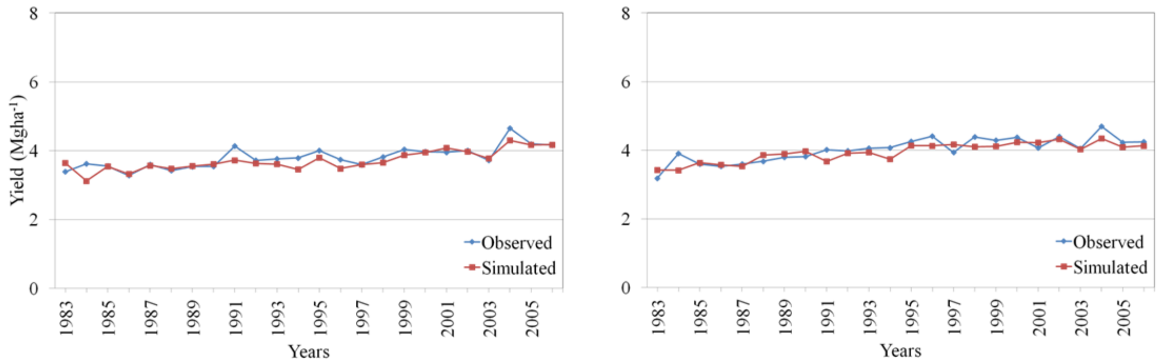

Figure 5. Temporal yield variability of observed (blue) and calibrated (red) winter wheat yields (Mg ha^{-1}) averaged over EU25 for the period from 1983 to 2006 considering (a) all climate zones and (b) only climate zones where more than 21 years of observed yield data were available for model calibration.

2.3.3 Calibration of other crops

Calibration results for other crops based on strategy 3 (phenology and growth parameters) were fairly satisfactory but some differences between observed and simulated yields were observed (Figure 6). Yield simulations were in better agreement with observations for winter crops wheat and barley as compared to the spring crops potato, maize and sugar beet yields, with the latter showing the largest differences. One reason for the larger differences between observed and simulated data for potato, sugar beet and maize as compared to the winter cereals could be the limited (or incorrect) availability of phenology data, particularly sowing dates. Discrepancies of some weeks between estimated and observed sowing date are more important and can have a large impact when simulating spring crops as compared to winter crops.

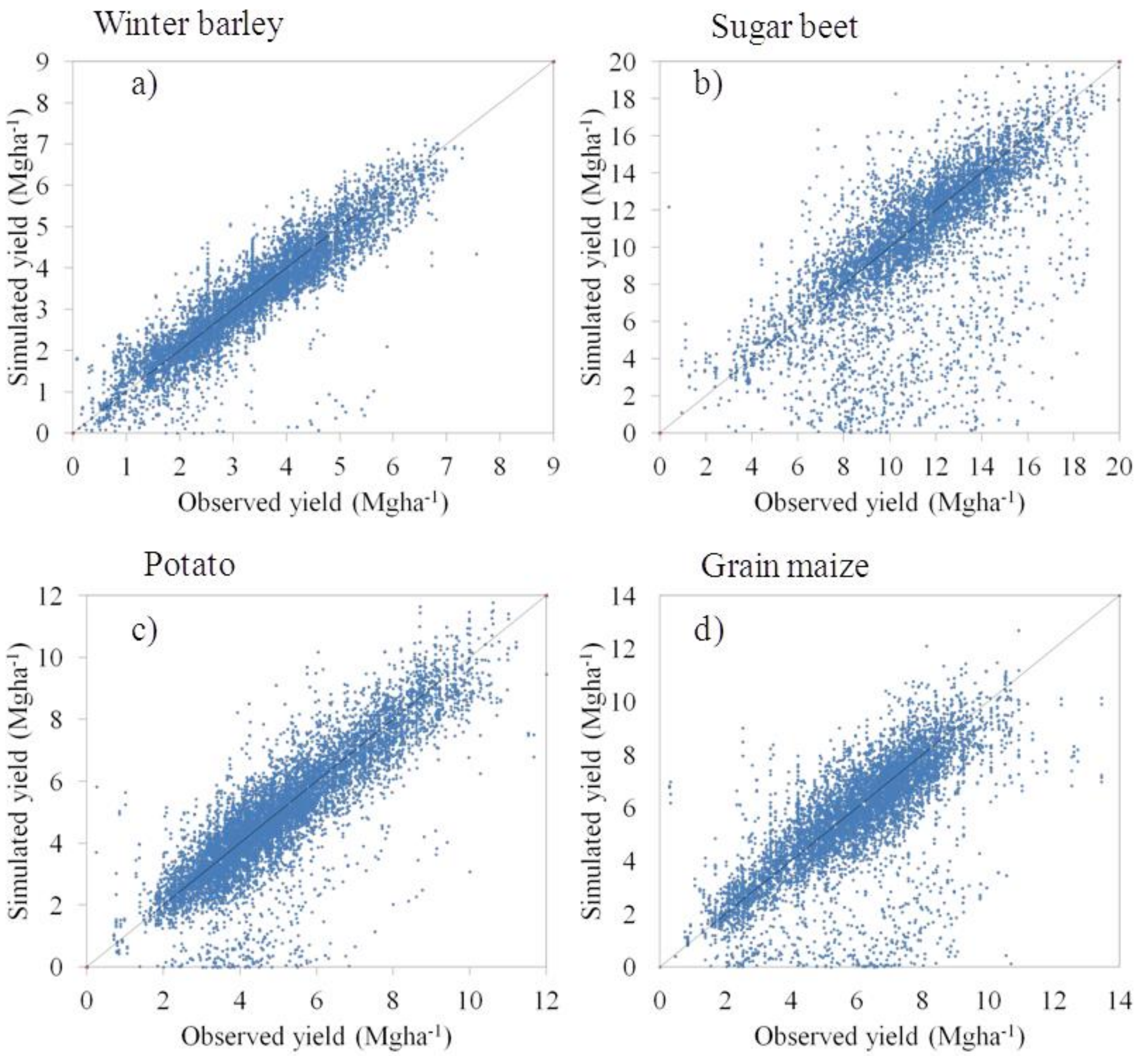

Figure 6. Comparison between observed and calibrated (strategy 3) yields of four crops in Europe (EU25) considering 533 climate zones and 24 years (1983 to 2006). (a) winter barley, (b) sugar beet, (c) potato, (d) grain maize (strategy 3).

Simulated spatial (Figure 7) and temporal (not shown) variability of yields in Europe for the selected crops are in acceptable agreement with observations. High productivity regions observed in Central and Western Europe (France, Netherlands, Belgium, Germany) for barley, potato and sugar beet are reproduced well as expected by the model. For grain maize, the highest yields are typically recorded in southern regions (Spain, Italy, and Greece) which the calibration strategy also captured. For some zones and crops, e.g. winter barley for Finland, phenological parameters were missing and no model calibration and simulation was performed.

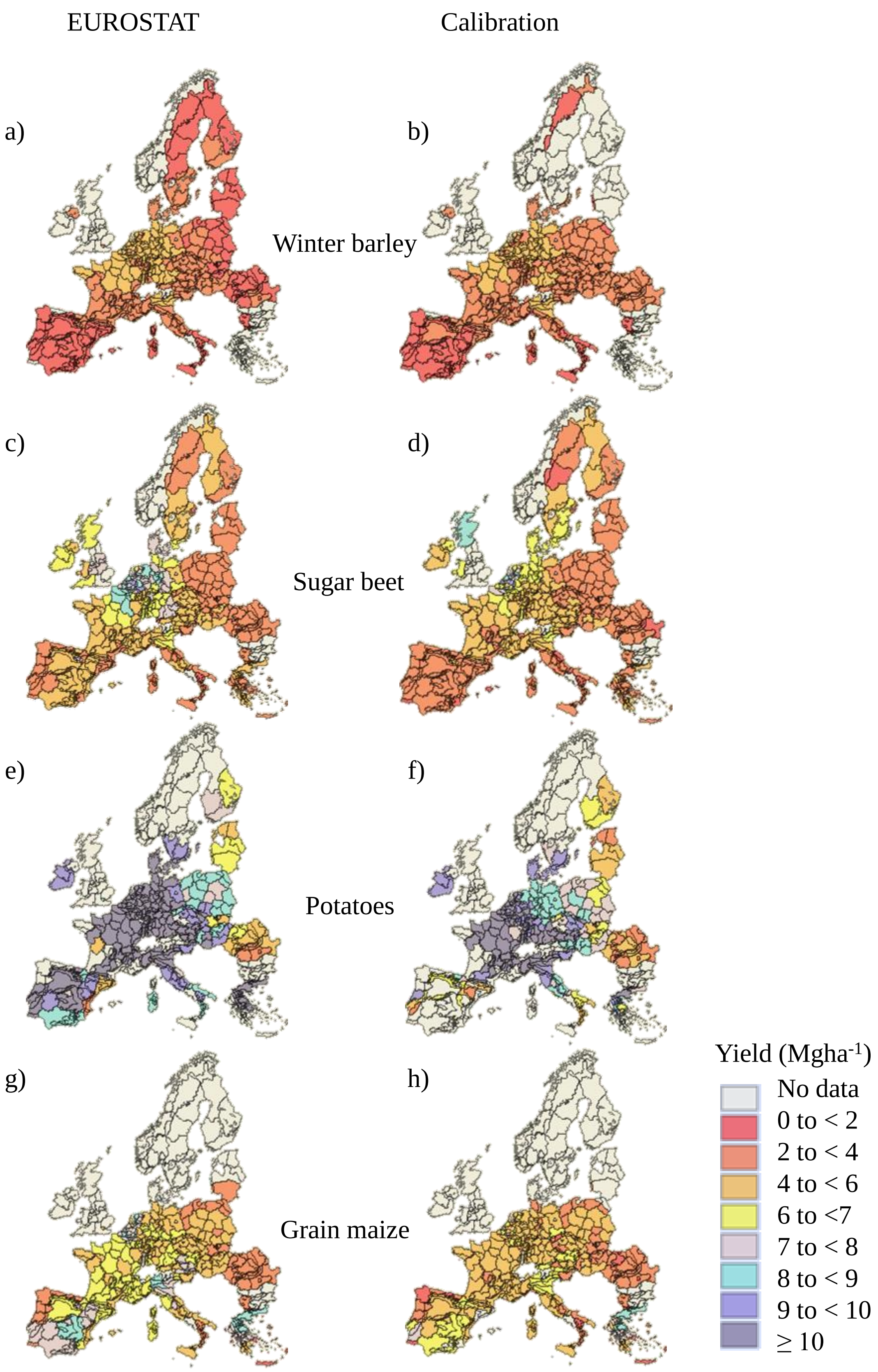

Figure 7. Spatial pattern of observed (a,c,e,g) and simulated (b,d,f,h) yields (Mg ha^{-1}) based on extended calibration of selected growth parameters (calibration strategy 3) for (a,b) winter barley, (c,d) potato, (e,f) sugar beet) and (g,h) maize in Europe averaged for the period 1983 to 2006.

2.4 Simulation of future yields

2.4.1.1 Impact of climate change

Climate change, without considering increasing atmospheric [CO_2] and advances in technology, causes a yield decrease for all crops and scenarios compared to the baseline yields (Figure 8a,d,g,j,m). The largest yield declines due to climate change were simulated with the GISS A1B scenario, a predominantly dry scenario (see section 2.2.6.1). However, differences between crops were observed. Projected climate change impacts on yields were largest for maize, approximately -1.7 Mg ha^{-1} (Figure 8m) and smallest for winter wheat, about -0.4 Mg ha^{-1} on average over EU25 (Figure 8a). We also realized that simulated responses to climate change were less for winter crops as compared to spring crops. This may be due to the longer vegetative period typical for winter crops, which allows winter crops to recover better from extreme events such as drought spells in spring. Also, climate change induced changes in growing season length due to temperature increase will be relatively smaller in winter as compared in spring crops.

2.4.1.2 Combined impacts of climatic change and increased [CO_2]

Taking into account elevated [CO_2] when simulating climate change impacts increases simulated yields for all crops and scenarios but with some variation. Yield increases are highest for the winter crops and compensate for the negative yield effect due to climate change (Figure 9b,e). In these crops projected future yields are higher than baseline yields for all scenarios. Also for the root crops, sugar beet and potatoes, the simulated yields are higher than the baseline yields in most scenarios, but for the scenario with the largest climate change impact, GISS A1B, the positive [CO_2] effect cannot compensate for the negative effect of climate change (Figure 8h,k). For grain maize there is almost no yield increase due to elevated [CO_2] (Figure 8n). Maize is a C4 plant and therefore elevated [CO_2] has no improving effect on radiation use efficiency but only on the transpiration rate (see section 2.2.1).

2.4.1.3 Combined impacts of climate change, increased [CO_2] and technology development

When both the effect of increased [CO_2] and technology development are taken into consideration together with the effect of climate change, simulated yield increases are considerable (Figure 8c,f,i,l) but with some noticeable differences among the crops. While for winter cereals and the root crops, yield increases are higher than the baseline for all future scenarios, simulated grain maize yields remain below the baseline yields (Figure 8o). Apparently, the simulated pronounced climate change effect on maize yield could not be compensated by increased [CO_2] and technology development. For the other crops, the highest yield increases are simulated for A1B scenarios (Figure 8c,f,i,l), in which [CO_2] abundance and temperature reach the highest values. Importantly, the consideration of technology development results also in larger differences of simulated yields among the scenarios.

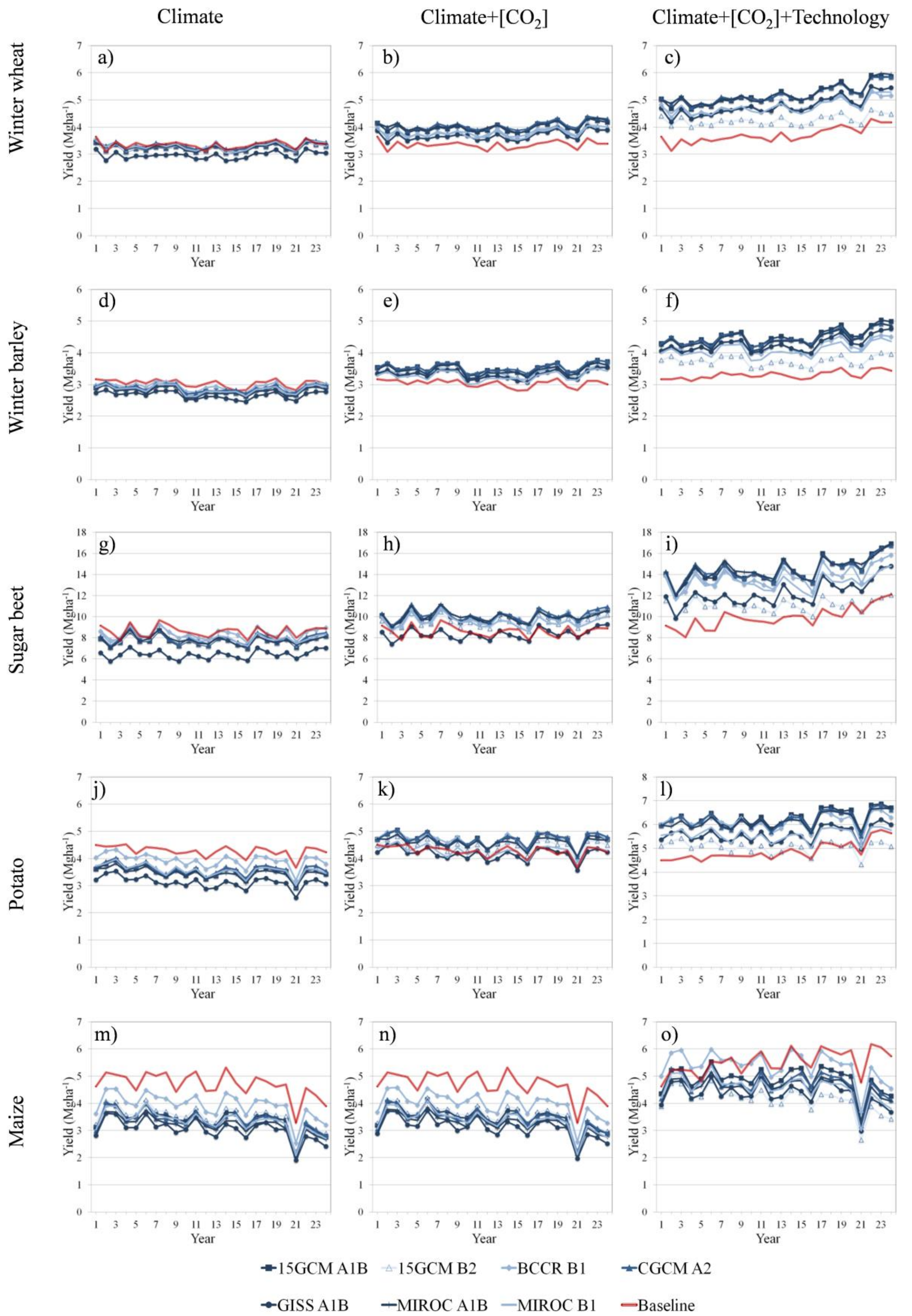

Figure 8. Simulated effects of (a,d,g,j,m) climate change, (b,e,h,k,n) climate change and increased [CO_2], and (c,f,i,l,o) climate change, increased [CO_2] and technological development on yields of five crops for 24 years in Europe (EU25) using four IPCC CC scenarios. Baseline and future scenarios are centred around 1990 and 2050 respectively. Crops considered are winter wheat (a,b,c), winter barley (d,e,f), sugar beet (g,h,i), potato (j,k,l) and maize (m,n,o).

An analysis of the spatial variability of simulated yields under combined changes in climate, [CO_2] and technology shows little differences among scenarios as can been seen from the comparison of yield simulation from A1B (15GCMs A1B) and B1 (MIROC B1), although some differences in the extent of yield changes in the individual regions can be noticed (Figure 9). For the winter cereals yield increases of 30% and more compared to the baseline are simulated for most regions. There are small areas on the Iberian and Italic peninsulas were yield decreases are projected compared to the baseline (Figure 9b,d). These declines are mainly due to the pronounced negative climate change effect which could not be compensated for by the positive [CO_2] and technology effect. The latter is relatively small due to the comparably small yield increases for these regions observed in the past. For potatoes and sugar beet yield increases are also simulated for most regions in Europe except for some areas in Southern Europe (Italy, Greece and Spain), and few regions in Poland and Finland, but in most of the cases the decreases do not surpass 10% in relation to baseline. For grain maize the spatial variability in yield changes ranges between <-30% to >30% (Figure 9i,j). Yield increases are highest in South-western Europe and yield declines are mainly projected for Eastern Europe.

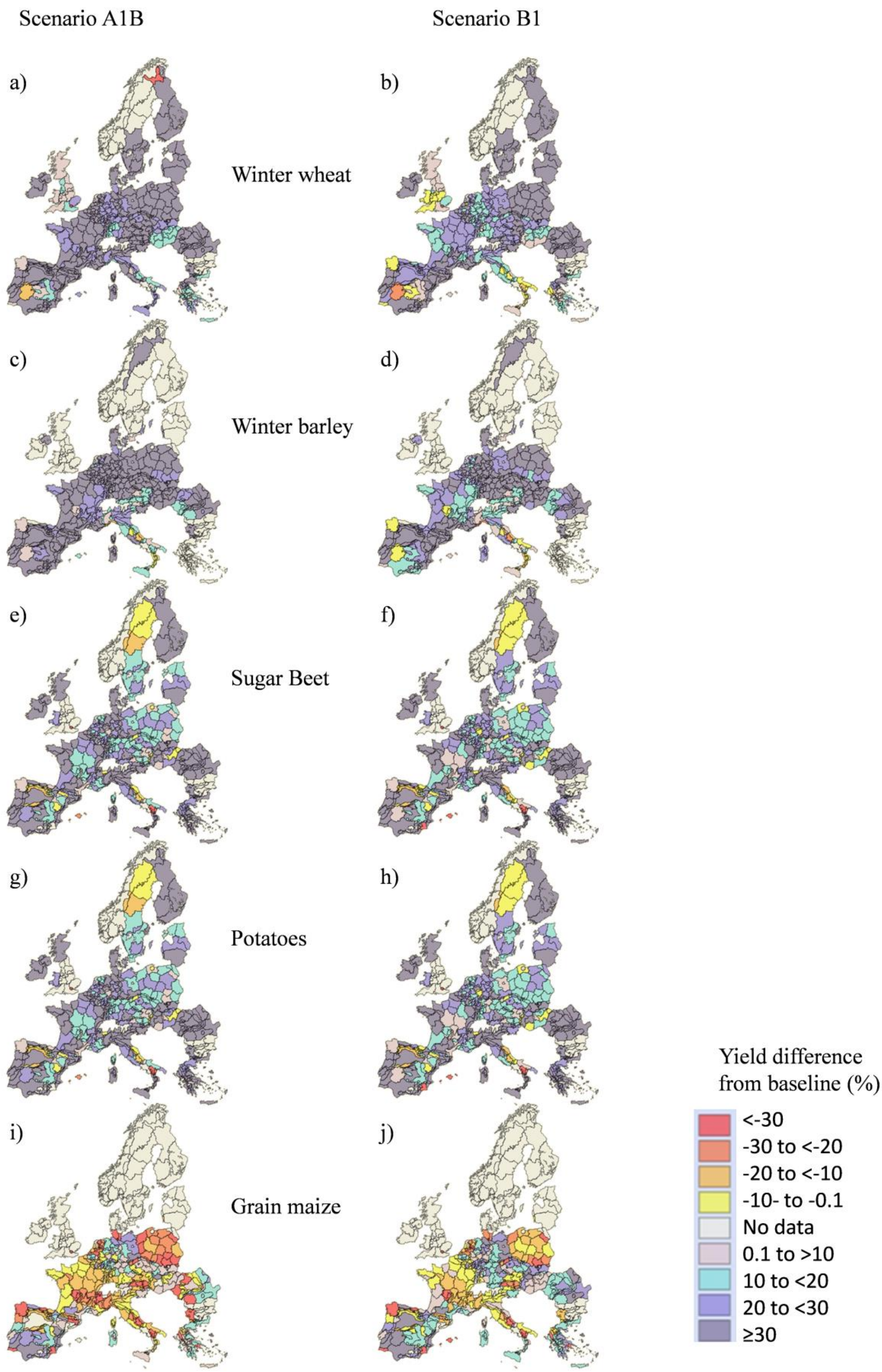

Figure 9. Differences between simulated baseline yields and yields from two climate change scenarios (a,c,e,g,i) A1B and (b,d,f,h,j) B1 for 5 crops over 24 years in Europe (EU25). The baseline and future time series are centred around 1990 and 2050, respectively. Crops considered are winter wheat (a,b), winter barley (c,d), sugar beet (e,f), potato (g,h) and maize (i,j).

Finally, we compared the temporal variability of our future projections with the baseline and the observed yield variability. Results are shown for three selected crops representing the range of responses for all five crops (Figure 10). The crop model LINTUL-FAST reproduces well the observed yield variability for all crops and most regions as was already described above (section 2.3.2). However, we identified some overestimations of the yield variability for potatoes and maize on the Iberian Peninsula. This may be due to an overestimation of the drought effect in the model. This overestimation can be expected for models applying the RUE concept instead of detailed photosynthesis routines (Rötter et al., 2012b). However, yield variability was reproduced satisfactorily in most regions.

There were only small changes in yield variability for the projected future scenarios for most crops, except for maize (Figure 10,h,i). The coefficients of variation (CV) of simulated grain maize decreased for the climate change scenarios as compared to the baseline on the Iberian Peninsula (Figure 10i). On the other hand, an increase in yield variability of maize due to climate change was observed for some regions in east Europe, mainly Poland (Figure 10i,).

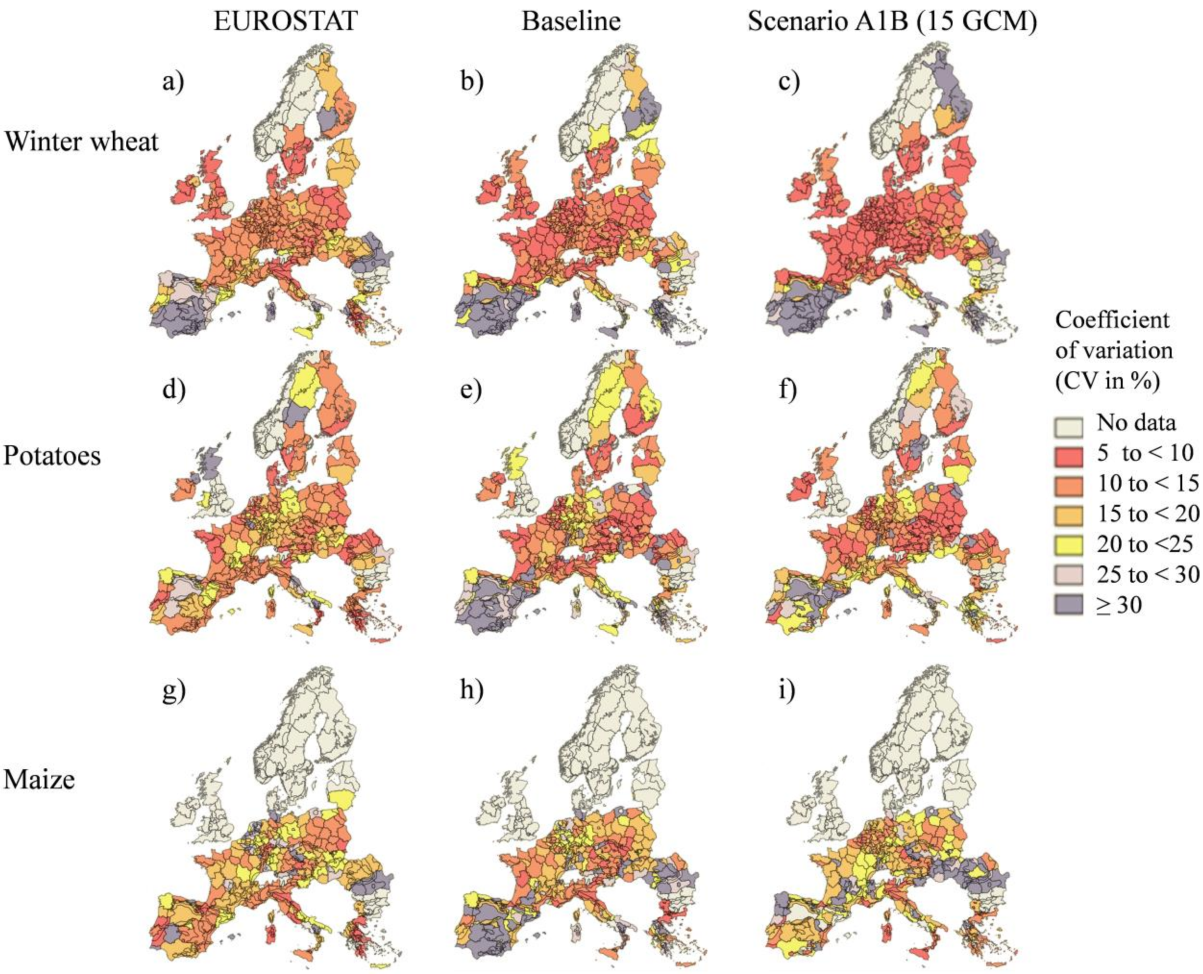

Figure 10. Coefficient of Variation (CV) over 24 years in Europe for (a,d,g) observed and simulated yields for (b,e,h) baseline (centred around 1990) and (c,f,i) the 15 GCM A1B scenario (centred around 2050). Crops shown are winter wheat (a,b,c), potato (d,e,f) and maize (g,h,i).

2.5 Discussion

2.5.1 Importance of model calibration

The present analysis is to our knowledge the first study in which the impacts of region-specific calibration of a crop model on yield simulations at continental scale (EU25) have been investigated. Our results are in agreement with earlier studies (Reidsma et al., 2009a; Therond et al., 2011) that consideration of differences in phenological development alone (strategy 1) does not suffice to capture variety differences among regions. Therond *et al.* (2011) proposed the use of a yield correction factor to account for regional yield differences not attributable to differences in phenological development. However, results have not been presented by the authors. In the present study we tested the use of such a correction factor and found a noticeable improvement of European-wide crop yield simulations when this factor was taken into account. However, further improvement on yield simulations was reached after a more extended calibration of selected growth parameters (strategy 3), (Figure 1). Furthermore, with this extended calibration strategy we were able to reproduce not only the spatial but also some of the temporal variability of crop yields (Figure 5). However, to which extent this confirms the capability of mechanistic crop growth simulation models to adequately capture the effects of climate on yield variability also at larger areas (Challinor et al., 2005; Hansen and Jones, 2000; Palosuo et al., 2011) needs further investigation as in our study model accuracy was only good in high yielding environments. Xiong *et al.* (2008) also improved simulations of yield variability in 16 Sub-Agro Ecological zones in China when considering region-specific model calibration. In their study the relative RMSE between simulated and observed yields, was from 15% to 74% after calibration. In our study the relative RMSE values were between 17% and 30% depending on the crop. Larger differences between observed and simulated data for potato, sugar beet and maize as compared to the winter cereals could be attributed to the limited (or incorrect) availability of phenology data, particularly sowing dates. Discrepancies of some weeks between estimated and observed sowing date are more important and can have a large impact when simulating spring crops as compared to winter crops.

The choice of the calibration strategy has implications for the reliability of model simulations (Challinor et al., 2009a). The results of the present study corroborate this affirmation and additionally provide a clear hint of the impact of the calibration strategy on the simulated effects of climate change on yield. Based on the present data we cannot assess which calibration strategy simulates climate change effects most accurately. Further research and most importantly data from independent regional time series will be needed to provide confirming evidence for this result.

Although our results indicate the importance of region-specific calibration of growth parameters, we are aware that the proposed strategy has some limitations. First, we have considered only three growth parameters of which we know that they refer to important growth processes such as light capturing, light conversion and effects of drought on biomass production. However, we have not tested that the choice of these parameters is sufficient as crop models typically comprise many more parameters. Such evaluation will require a larger effort and was not the aim of this study. We have also not tested whether the calculated parameter values correctly represent the varieties grown in a specific region. As our calibration is based on observed yields from regional statistics, we cannot exclude that other effects related to factors such as pests and diseases or limitations of nutrients have affected these observed yields and thus our calibrated parameters. Such testing would require location specific information about crop growth and development processes for which European-wide data is not available. As crop models are typically calibrated using location specific information of crops grown on a sample of small plots (Faivre et al., 2004), some effort is needed to better understand the up-scaling of these parameters from the plot to region scale.

Therefore, without further investigations we cannot recommend the calibration of growth parameters on regional yield statistics for large scale impact assessment. Although results of the present study suggest some improved model behaviour if growth parameters are calibrated, the use of a yield correction factor is still more meaningful. However, multiplication of simulated yield with a yield correction factor may also affect the yield variability resulting in larger RMSEs and therefore more detailed investigations are required to better clarify this effect.

In our study we have also not investigated the potential of other calibration approaches such as the Bayesian approach which is increasingly used also in crop and ecosystem modelling (Lehuger et al., 2009; Reinds et al., 2008; Tao et al., 2009; Tremblay and Wallach, 2004). By applying the theorem of the conditional probability, the Bayesian approach utilizes output variables, for example yield, to calculate a posterior calibrated parameter values distribution based on a prior probability distribution which is given by the quantified uncertainty of the parameter values of a model (Van Oijen et al., 2005). This approach may provide a more comprehensive overview about the relative importance of parameters capturing regional differences in crop growth and yield. It could also give more qualified information about the parameter uncertainty and insides to appropriate parameter-space sampling. However, with our study we could show that uncertainty due to restricted parameterisation can be large for both simulated yields and climate change effects on yields and that model calibration of growth parameters for individual regions can substantially improve model accuracy as compared to the use of one general set of growth parameters (e.g. De Wit et al., 2010).

2.5.2 Impacts of climate change, [CO_2] increase and technology development

Our results suggest that for EU25 the negative effects of climate change on crop yields range between 12% and 34% depending on the crop and region. Climate change effects are less pronounced for winter cereals (barley and wheat) as compared to tuber crops (potatoes and sugar beet) or other spring crops (maize). One possible explanation, still subject of further investigation, is the longer vegetative period for winter crops which may allow the winter crops to better cope with extreme events such as drought spells in spring. Also, changes in growing season length due to temperature increase will be relatively smaller in winter as compared to spring crops.

The simulations with the driest scenario GISS A1B resulted in the strongest negative influence on yields even when taking the [CO_2] fertilization effect (Rötter and Van De Geijn, 1999; Tubiello et al., 2007) into account. The overall range in simulated yield changes among scenarios is large but differed among crops. Again, the range was less pronounced for winter as compared to spring crops. For the latter, on average for EU25 the differences among scenarios were larger than the climate change effect within one scenario or the simulated temporal yield variability.

These simulated changes are more pronounced than the projection by Ewert *et al.* (2005) who calculated a climate change effect by 2050 which was on average over 15 EU member countries less than 3% yield reduction. Such results point to the tendency of crop simulation models to project higher effects of climate changes than statistical approaches. This may be explained by the fact that crop models primarily consider the effects of climate factors on crop growth and development. Effects of other factors such as weeds, pests and diseases are generally not considered explicitly by crop models based on mechanistic modelling but on statistical-empirical approaches (Savary et al., 2006). Most often the influence of such factors is expressed by a yield reduction factor as in the case of GLAM (Challinor et al., 2004). Large scale evaluation of crop models is also in an early stage (Van Oijen and Ewert, 1999). Again, experimental data will be required to support such evaluation.

Effects of elevated atmospheric [CO_2] enhanced yields mainly for C3 crops to an extent which is consistent with data from FACE experiments (Ainsworth and Long, 2004; Long, 2006; Manderscheid and Weigel, 2007). Increasing [CO_2] concentration stimulated yields in wheat, barley, sugar beet and potatoes by 14%, 11%, 14% and 7% respectively, with small differences between years and regions.

However, most substantial yield changes were projected when considering the effect of technology development, which is consistent with earlier results (Ewert et al., 2005). Importantly, considering a technology effect not only increased the crop yields but also increased the differences between the scenarios. Projected yields were highest for the scenarios CGCM A2 and 15GCM A1B and smallest for the scenario 15GCM B2. This is due to the assumptions of scenario family A (Nakicenovic and Swart, 2000) in which higher intensification and thus a more advanced technology development is considered.

Clearly, considering the effects of climate change, atmospheric [CO_2] elevation and technology development separately had two main implications for our yield projections. On the one hand, the yield decreasing effect of climate change was compensated and partially superseded when atmospheric [CO_2] elevation and technology development were taken into account which is in good agreement with earlier research (Ewert et al., 2005). On the other hand, the yield differences between scenarios become greater when considering atmospheric [CO_2] elevation and technology development.

Finally, our results show some changes in yield variability under climate change (Figure 10). However, these changes were mainly observed for maize and differed considerably depending on the region from decreasing to increasing yield variability under climate change. Other studies have reported increased yield variability as an impact of climate change in Europe (Iglesias et al., 2010; Jones et al., 2003b; Porter and Semenov, 2005). However, in the present study we have not considered an approach to model the effects of extreme temperature stress (Asseng et al., 2011; Porter and Gawith, 1999; Porter and Semenov, 2005). Modelling such effect is likely to result in a more pronounced yield variability under climate change, as it has been recently shown in a global assessment for four crops (Teixeira et al., 2013).

2.6 Conclusions

The present study investigated the importance of crop model calibration to enhance assessment of climate change impacts on crop yield at regional scale. We find that considering regional differences of model parameters related to crop growth in addition to crop phenology can considerably improve yield simulations at continental scale (EU25). Calibration also affects simulations of climate change impacts on yields. These results suggest that regional projections with crop models can be improved if they are calibrated with region-specific data. However, proper calibration of crop growth and development parameters requires data which are presently not sufficiently available for entire Europe. Our results also confirm earlier studies about the importance of considering not only the effects of changes in weather variables, but also increased atmospheric [CO_2] and technology development for future yield estimations. Particularly, consideration of technology development can have substantial impacts on yield projections. Further investigation is required to reduce uncertainty in the assumptions regarding technology development. The considered crops respond differently to climate change which also calls for extending climate change studies to a larger range of crops. The considered ensemble of climate change scenarios results in a range of yield responses which again is more pronounced when technology development is considered. As some of this technology development refers to yield improvements, future research on improving model calibration for large scale climate change studies will also need to address temporal changes in model parameters.

Chapter 3

Characteristic ’fingerprints’ of crop model responses to weather input data at different spatial resolutions

The present chapter has been published as:

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. doi 10.1016/j.eja.2013.04.003

3. Characteristic 'fingerprints' of crop model responses to weather input data at different spatial resolutions

3.1 Introduction

Process-based crop growth models (further on referred to as crop models) are increasingly being utilized as tools for assessing the regional impact of climate variability and change on crop production (Challinor et al., 2009b; De Wit et al., 2010; Hansen and Jones, 2000; Jagtap and Jones, 2002; Reidsma et al., 2009a; Rötter et al., 2011b; Therond et al., 2011; Tubiello and Ewert, 2002; White et al., 2011). However, the regional applicability of crop models is critically discussed for two main reasons. First, crop models have typically been developed and validated at the field scale (Boons-Prins et al., 1993; Brisson et al., 1998; Stockle et al., 2003; Van Ittersum et al., 2003b; Williams et al., 1983) and scale-change issues emerge when applying crop models at larger spatial extent, e.g. regions (Ewert et al., 2011b). Second, crop models require environmental (weather and soil) and agricultural management input data that are seldom available for larger areas at the required level of detail (Ewert et al., 2011b; Faivre et al., 2004; Leenhardt et al., 2006).

Accurate weather input data are crucial to obtain coherent yield simulations when investigating the effects of climate change and variability on crop yields in larger regions (Hansen and Jones, 2000). Since weather data are measured only at a limited number of meteorological stations within a region, it is necessary to estimate the values of the required weather variables for the appropriate simulation-scale (Faivre et al., 2004). The uncertainty introduced through such estimations is largely unknown but should be reported when simulating crop yields. It has been found that even the estimation methods yielding the lowest bias in comparison to measured daily solar radiation generate random errors when simulating biomass production with the models DSSAT-CSM and WOFOST (Trnka et al., 2007). Additionally, a common practice when applying crop models regionally is to use weather input data spatially interpolated onto grid cells of various resolutions (e.g. De Wit et al., 2005; Mearns et al., 2001; Van Bussel et al., 2011a). A grid cell consists of a multiple set of weather parameters derived from interpolation of weather station data assigned to single area units or cells (e.g., Venäläinen et al., 2005). The size and boundaries (spatial distribution) chosen to build the individual cells inevitably causes a biasing error in relation to the measured data and might consequently impact negatively the validity/accuracy of the spatialized weather data. The mentioned biasing error which might lead to ecological fallacy is addressed in the literature as the modifiable areal unit problem (MAUP) (Dark and Bram, 2007; Holt et al., 1996; Hui, 2009; Unwin, 1996). The MAUP emerging from spatialization of weather represents an additional source of uncertainty and should be taken into consideration when analysing the results of regional crop model applications (Holt et al., 1996; Unwin, 1996)

The effect of the spatial resolution of weather data has been assessed for countries in Europe (Germany and France). Precipitation and radiation data from General Circulation Models (GCM) which were down-scaled to a resolution of 50 km x 50 km were more appropriate for forecasting national yields than data from coarser spatial resolutions (De Wit et al., 2005). Likewise, Mearns *et al*. (2001) reported effects on yield simulations by two different spatial resolutions of climate scenario data, one from a regional climate model (RCM) and another from a GCM, and concluded that climate scenarios with higher resolution are more suitable for climate impact assessment. Recognizing the importance of weather input data resolution, Van Bussel *et al*. (2011a) compared the effect of spatial aggregation of weather and emergence dates on phenological stages simulated by the model AFRCWHEAT2. According to their findings for winter wheat in Germany, cells with a maximum area of 100 km × 100 km can be considered a sufficiently appropriate resolution to simulate the length of the growing season. Model uncertainty also increases with the temporal aggregation of input data, It has been shown for European conditions that the temporal aggregation of weather data causes overestimation of simulated yields which, however, depends on the detail of the modelling approach used (Van Bussel et al., 2011b).

Although the studies cited above underline the importance of weather input data resolution on crop model simulations, the effects of spatially aggregating weather input data on simulated yields have only been partially assessed until now. In a national scale study, Olesen *et al*. (2000) simulated winter wheat yields in Denmark considering the effect of aggregating weather (without precipitation) and soil data at 1 km x 1 km and 10 km x 10 km resolutions, and compared model output to aggregated observed yields. They found that spatially detailed weather information is not necessary for a whole-country assessment; nevertheless, increased spatial resolution of weather data might be required for assessing productivity of Danish sub-regions as has also been demonstrated for Germany (Nendel et al., 2013).

In addition to the uncertainty associated with different spatial and temporal scales in crop model applications, there is an increasing interest in comparing crop models and quantifying the model related uncertainty (Palosuo et al., 2011; Rötter et al., 2012b). To date, the possible interaction between modelling approach and spatial resolution of weather input data has not been addressed yet (Rötter et al., 2011a). Even less is known about the degree of uncertainty in the observed yield data used for model calibration, validation and evaluation purposes. Only few studies consider the spatial or temporal distribution of simulated yields as compared to the distribution of observed yields (e.g. Easterling *et al*., 2007). In response to these knowledge gaps, this study systematically assesses the effects of changes in weather data resolution on simulated regional yields of four crop models. More specifically, we evaluate the influence of spatially aggregated weather input data on the frequency distributions of simulated yields in a region, in addition to commonly used centred statistics of means, medians and ranges. We compare the frequency distributions of simulated yields for different models clarifying whether a model-specific distribution pattern, a so-called “fingerprint”, can be identified, and to which extent it changes with the spatial resolution of input data. Finally, we assess the effects of aggregating model inputs *versus* model outputs on the distribution of simulated yields, as well as the effects of aggregating observed yields in comparison to the aggregation of simulated yields.

3.2 Materials and Methods

3.2.1 Models

The models utilized in this study were: LINTUL-SLIM(Addiscott and Whitmore, 1991a; Angulo et al., 2013a,), DSSAT-CSM(Jones et al., 2003a), EPIC (Gassman et al., 2004), and WOFOST (Boogaard et al., 1998; Van Diepen et al., 1989). These models use different approaches for simulating plant growth and development and differ in complexity regarding how they describe physiological processes and sub-systems (see, e.g. Palosuo et al., 2011; Rötter et al., 2012a). All models have already been applied and calibrated for Finnish conditions (Angulo et al., 2013a; Palosuo et al., 2011; Rötter et al., 2012a; Salo *et al*. in preparation). The RMSE values obtained by Rötter et al. (2012a) for simulated spring barley in Jokioinen were 1.31 Mg ha^{-1} for LINTUL-SLIM, 1.5 Mg ha^{-1} for DSSAT-CSM, and 1.98 Mg ha^{-1} for WOFOST. For EPIC the RMSE value was 0.77 Mg ha^{-1} (Salo *et al*. in preparation).

A summarized description considering the major crop growth processes relevant for the present study based on the description by Palosuo *et al*. (2011) can be found in Table 3.

Regarding the level of detail of the processes of light interception and light utilization (for more details see Adam et al., 2011), the models use three different approaches. LINTUL-SLIM and DSSAT-CSM use a detailed approach for simulating leaf area index (LAI) dynamics, based on temperature and leaf dry matter supply, driven by the development stage of the crop (Spitters, 1990). They combine this detailed LAI approach with a simplified approach for estimating biomass production, utilizing the radiation use efficiency (RUE) concept (Monteith and Moss, 1977). EPIC utilizes a simplified approach for simulating LAI dynamics based on a forcing function (Williams et al., 1983), combined with the simplified RUE approach. Finally, WOFOST applies a detailed approach for simulating LAI dynamics combined with a detailed approach for biomass production based on the description of the photosynthesis and respiration to describe the production of biomass (Van Ittersum et al., 2003b).

All models describe crop development stage as a function of temperature and photoperiod (Slafer and Rawson, 1996; Van Ittersum et al., 2003a). In LINTUL-SLIM and WOFOST, final grain yield is calculated as a function of total daily dry matter allocation to different plant organs according partitioning functions depending on crop development stage (Van Ittersum et al., 2003b). Grain yield in DSSAT-CSM is calculated from simulated grain number per ear determined by the estimated biomass accumulation during a fixed thermal time phase before flowering, grain weight depending on the length of grain filling period, and ear number per area unit (Langensiepen et al., 2008). In EPIC final grain yield is calculated as a function of total biomass and harvest index, given as an input parameter (Mearns et al., 1999).

Table 3. Major processes determining crop growth and development of the models applied in this study (Modified from Palosuo et al., 2011).

Model	LINTUL-SLIM	DSSAT-CSM	EPIC	WOFOST
Version	220	4.0.1.0	0509	7.1
Leaf area development and light interception [a]	D	D	S	D
Light Utilization [b]	RUE	RUE	RUE	P-R
Yield formation [c]	Y(Prt)	Y((GnGw,En.Prt)B)	Y(HI,B)	Y(Prt,B)
Crop Phenology [d]	f(T,DL,V)	f(T,DL,V)	f(T,DL)	f(T,DL)
Stresses involved [e]	W	W,N	W,N,P	W
Water dynamic [f]	C	C	C	C
Evapo-transpiration [g]	PM	PT	PM	P
Soil CN-model [h]	-	CN,P(4),B	CN,P(4),B	-

[a] Leaf eara development and light interception: S = simple or D = detailed approach.

[b] Light utilization or biomass growth: RUE = Simple (descriptive) Radiation use efficiency approach, P-R = Detailed (explanatory) Gross photosynthesis-respiration.

[c] Y(x) Yield formation depending on: HI = fixed harvest index, B = total (above ground) biomass, Gn = number of grains, Prt = partitioning during reproductive stages, Gw=grain weight, En=earn number.

[d] Crop phenology is a function (f) of: T = temperature, DL = photoperiod (day length), V = vernalisation; O = other water/nutrient stresses effects considered.

[e] Stresses involved: W = water, N = nitrogen stress, P = phosphorus stress.

[f] Water dynamics approach: C = capacity approach, R = Richards approach.

[g] Method to calculate evapo-transpiration: P = Penman, PM = Penman-Monteith, PT = Priestley-Tailor

[h]Soil-CN model, N = nitrogen model, P(x) = x number of organic matter pools, B = microbial biomass pool.

All models contain modules considering plant-soil-water dynamics. LINTUL-SLIM calculates potential evapotranspiration using the Penman-Monteith equation according to Allen *et al,* (1998b). In WOFOST potential evapotranspiration is calculated with the Penman formula (Penman, 1956), adapted according to Frère and Popov (1979). DSSAT-CSM applies the Priestly-Taylor equation (Priestley and Taylor, 1972) and EPIC, the Penman-Monteith equation (Monteith and Greenwood, 1986).

3.2.2 Study region and model input

The study area is located in South-western Finland, where spring barley, *Hordeum vulgare*, is widely cultivated. The considered region is 400 km^2 in size, of which approximately 230 km^2 are part of the Yläneenjoki river catchment. The area was chosen because of the large amount of available data on yield from farmers' fields and for weather data. According to the environmental stratification of Europe (Metzger et al., 2005), the area falls into the boreal environmental zone. The climatic characteristics of the Jokioinen experimental station located near the south-western border of the study area are presented in Table 4.

Table 4. Characteristics of Jokioinen experimental station: longitude, latitude, altitude and long-term agro-climatic conditions from 1971 to 2000. (Adapted from **Rötter et al., 2012a**).

Characteristic		Value	
Longitude		23°30'E	
Latitude		60°48'N	
Altitude (m.a.s.l.)		104	
Mean annual temperature (°C)		4.3	
Mean annual precipitation (°C)		506	
	Lowest	Mean	Highest
Mean temperature May-August (◦C)	11.5	13.4	15.0
Sum of temperatures above 0◦C May-August (◦C)	1411	1644	1848
Sum of precipitation May-August (mm)	143	252	360
Shortwave radiation flux (Wm^{-2})	171	205	229

Weather data

Weather data for the period from 1994 to 2005 were obtained from two sources: the basic climate data set provided by the Finnish Meteorological Institute with a 10 km × 10 km grid cell resolution (Venäläinen et al., 2005) and the recordings from the Jokioinen meteorological weather station (Drebs et al., 2002). Data included daily measurements of global solar radiation (MJ m^{-2} d^{-1}), maximum air temperature (°C), minimum air temperature (°C), rainfall amount (mm d^{-1}) and vapour pressure (hPa). Observed daily values of wind speed (m s^{-1}) from the Jokioinen station were transmitted to all considered grid cells and aggregated grid cell as they were not available in the gridded data set.

Soil data

A clay-loam soil, typical for the region, was used as standard soil for all simulations. Originally, a heavy-clay soil type was also considered in the simulations. However, a preliminary analysis revealed that this soil did not affect the aggregation results (not shown) so that it was ignored for further analysis. For the clay-loam, profile-average water content values at field capacity and wilting point were 0.425 and 0.259 (m^3m^{-3}), respectively. As no observed values of initial soil water content at the beginning of the simulations runs were available, simulations of initial soil water content from EPIC were used as input values for all models. To obtain these values from EPIC, the model was initialized utilizing four additional years of weather data prior to 1994, i.e. from 1990 to 1993. For each year of the period from 1994 to 2005 the soil water content simulated by EPIC for each corresponding sowing date in a year was extracted from the daily water balance and used as initial soil water content for the three remaining models.

Crop data

Information on observed sowing dates and yields for spring barley at the Yläneenjoki river catchment was available from the Finnish Study of Monitoring the Impacts of Agri-environmental Support Scheme (MYTVAS). The MYTVAS-database is reported and summarized in Palva *et al.* (2001), Pyykkönen *et al.* (2004), Mattila *et al.* (2007) and Turtola and Lemola (2006). It provides information from 400 to 600 parcels on farms within the study region. The sowing dates corresponding to the values of the median, 25 and 75 percentile of the sowing dates distributions for the period from 1994 to 2005 were considered as model input data.

Two cultivars were selected for the study region, Annabell and Scarlett. While Annabell is relatively late maturing (approx. 97 days after sowing), Scarlett represents the average maturity type of modern barley cultivars in Finland (approx. 93 days after sowing) (Hakala et al., 2011).

3.2.3 Set-up of simulation study

Spatial resolutions of weather data

A summary of the simulation steps is given in Table 5. In total, seven simulation steps were performed. For each step identical data on soil type, sowing dates and barley varieties (see section 2.2) were used. The difference among simulation steps refers to the different spatial weather data resolutions. Simulations for the three different sowing dates and two varieties were considered to generate a representative variability in grain yields for the area; their effects on yield were not specifically analysed as it was not the aim of the study. The weather data processing for each simulation step was done as follows:

Table 5. Set-up of the comparison study for each model

Step	Weather resolution Cells		Number of simulated yields (V × Sd × Cells × 12 years)*
1	1	Weather station	72
2	4	10x10km grid cells	288
3	1	20x20km grid cell	72
4	1	50x50km grid cell	72
5	1	100x100km grid cell	72
6	25	50x50km grid cells	1800
7	100	100x100km grid cells	7200

*V = 2 varieties (Annabell and Scarlett); Sd = 3 sowing dates; Cells = weather support unit; 12 years = considered period from 1994 to 2005.

Step 1.- The models were run using the weather data obtained from the Jokioinen meteorological station as input (Figure 11).

Step2.- Weather data from the 10 km × 10 km grid provided by the Finnish Meteorological Institute were used as inputs for the yield simulations. As the study area extends over mainly four (10 km × 10 km) grid cells (no. 453,454,500,501, see Figure 11), weather data of these four cells were used as input for yield simulations.

Step 3.- Weather data of the four grid cells used in Step 2 were averaged for every variable and day, obtaining an aggregated grid cell of 20 km × 20 km, which was used as weather input for yield simulations.

Step 4.- As in step 3, weather data of 25 (10 km × 10 km) grid cells (no. 357-361, 404-408, 451-455, 498-502, 545-549 see Figure 11), were averaged for every variable and day, obtaining an aggregated grid cell with a resolution of 50 km × 50 km which was used as input data for yield simulations.

Step 5.- As in step 3, weather data of 100 (10 km × 10 km) grid cells (no. 260-269, 308-317, 355-364, 402-411, 449-458, 496-505, 543-552, 589-598, 636-645, 683-692 see Figure 11) were averaged and the resulting aggregated 100 km ×100 km grid cell was used as input data for yield simulations.

Step 6.- Differently from steps 1 to 5, yields were simulated individually for each individual 10 km × 10 km grid cell (a total of 25) within the 50 km × 50 km mega-cell used in step 4, 25 in total (see b in Figure 11). The simulated yields were subsequently aggregated to the correspondent 50 km × 50 km mega-cell.

Step 7.- As in step 6, yields were simulated individually for each individual 10 km ×10 km grid cells (a total of 100) within the 100 km ×100 km mega-cell used in step 5 (see c in Figure 11). The simulated yields were subsequently aggregated to the correspondent 100 km × 100 km mega-cell.

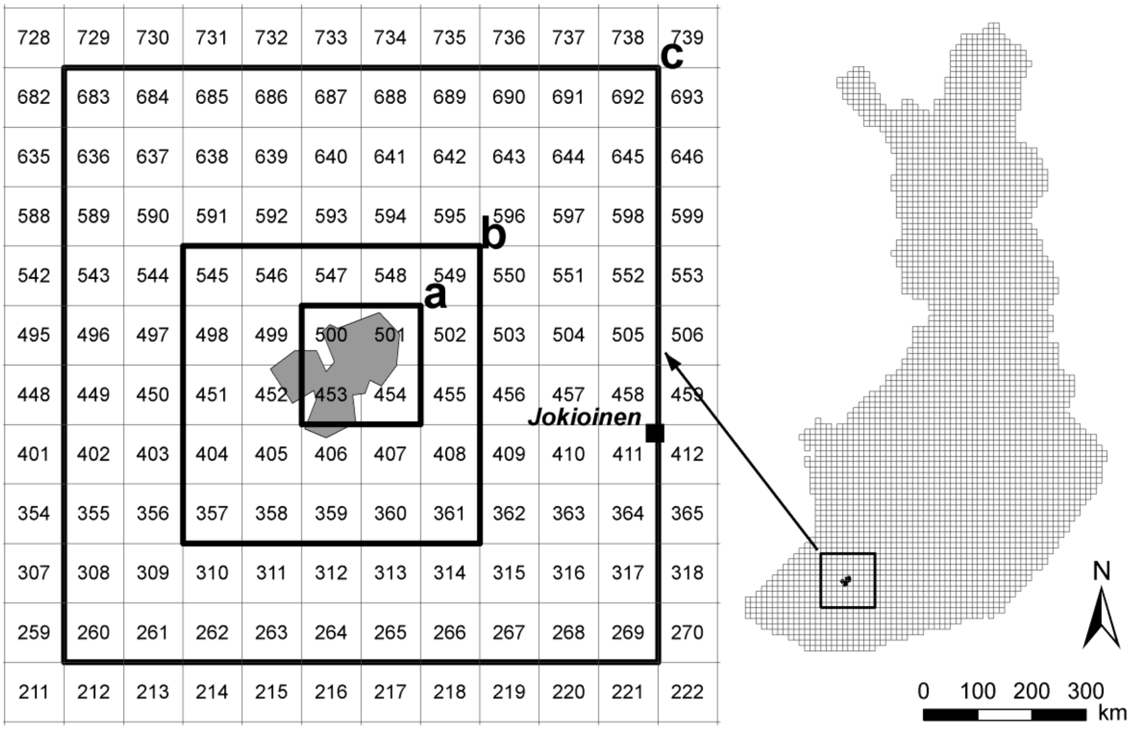

Figure 11. Region and aggregation units considered in the study. The number in each cell marks the identifying number of each grid cell. Weather data were available with 10 km × 10 km resolution.- Other aggregation units tested refer to: a, 20 km × 20 km; b, 50 km × 50 km and c, 100 km × 100 km). The location of the Jokioinen weather station is marked with a box. Sites with observed yields are located within the shaded area.

Steps 1 to 5 were performed to test the influence of the spatial resolution of weather data on the simulated yields including their distributions for all considered models. Steps 6 and 7 were performed only with the model LINTUL-SLIM, to evaluate the effect of the aggregation strategy, i.e. the aggregation of weather input data as compared to the aggregation of simulated yields. Figure 12 illustrates the two aggregation strategies exemplifying a resolution of 20 km × 20 km. For the first strategy (Figure 12a), yields were simulated using aggregated weather data (steps 3, 4, 5). For the second strategy, simulated yields obtained in steps 2, 6 and 7 were averaged for every year×cultivar×sowing date combination to obtain an aggregated yield distribution for a grid cell resolution of 20 km × 20km, 50 km × 50 km and 100 km × 100 km, respectively (Figure 12b).

Effect of weather data resolution on simulated yield distributions

Simulated yields were analysed with respect to the influence of both weather data resolution and model on the distribution of simulated yields. Results are presented in the form of bean plots. Similar to box and whisker plots, bean plots present the range of the data sample without further assumptions of the distribution and the median. In addition, bean plots show a density trace (contour line of the bean) of the analyzed data providing information about the frequency distribution of the data. In the present study, the normal (Gaussian) kernel was used for the calculation of the density trace. For a detailed description of the bean plot implementations see Kampstra (2008).

Effect of spatial aggregation on observed yields

Finally, we analysed the extent to which the spatial aggregation of observed yields affected the observed frequency distributions and introduced uncertainty into the model evaluation. In total, 6300 yield observations were available from 12 years and more than 400 locations spreading across four 10 km ×10 km grid cells (Figure 11). Distributions of observed yields were evaluated for individual site data and two different aggregations, i.e. 10 × 10 km and 20 km × 20 km. The distribution of yields for the 10 km × 10 km resolution considered the weighted averages of yearly measured yields for the four corresponding weather grid cells separately (48 values). The distribution of yields at 20 km × 20 km resolution considered the weighted averages of yearly measured yields of all four weather grid cells (12 values).

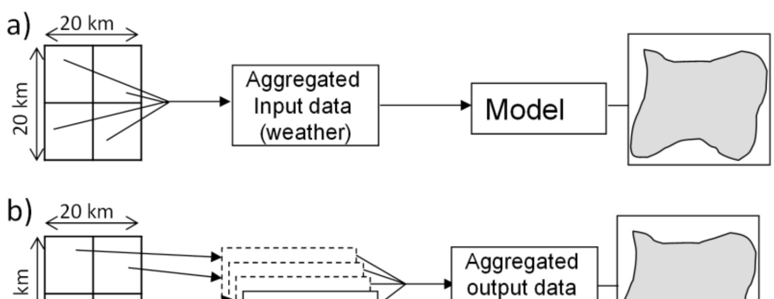

Figure 12. Schematic representation of scaling methods compared in this study, referring to: a, aggregation of weather input data and b, aggregation of outputs (adapted from Ewert et al., 2011b).

3.3 Results

3.3.1 Influence of weather data resolution on simulated yield distributions

Consistently for all four models, the simulated yields and their distributions were hardly affected by the five weather resolutions. The most noticeable effect of a resolution-change on the distributions of simulated yields was the transition between using point data (from one individual weather station) to aggregated (grid-based) data (Figure 13). For example, the median of LINTUL-SLIM yields simulated with the single station data was 3.70 Mg ha^{-1}, whereas the median of simulated yields using aggregated weather data was higher but ranged only between 4.44 and 4.70 Mg ha^{-1} depending on the resolutions considered (**Table 4**). The median of simulated yields using point data was also smaller for the models EPIC and WOFOST compared to the median of yields using aggregated data (Figure 13). In contrast, for DSSAT-CSM (Figure 13) the distribution of yields simulated with the weather station data showed the highest median value (**Table 4**).

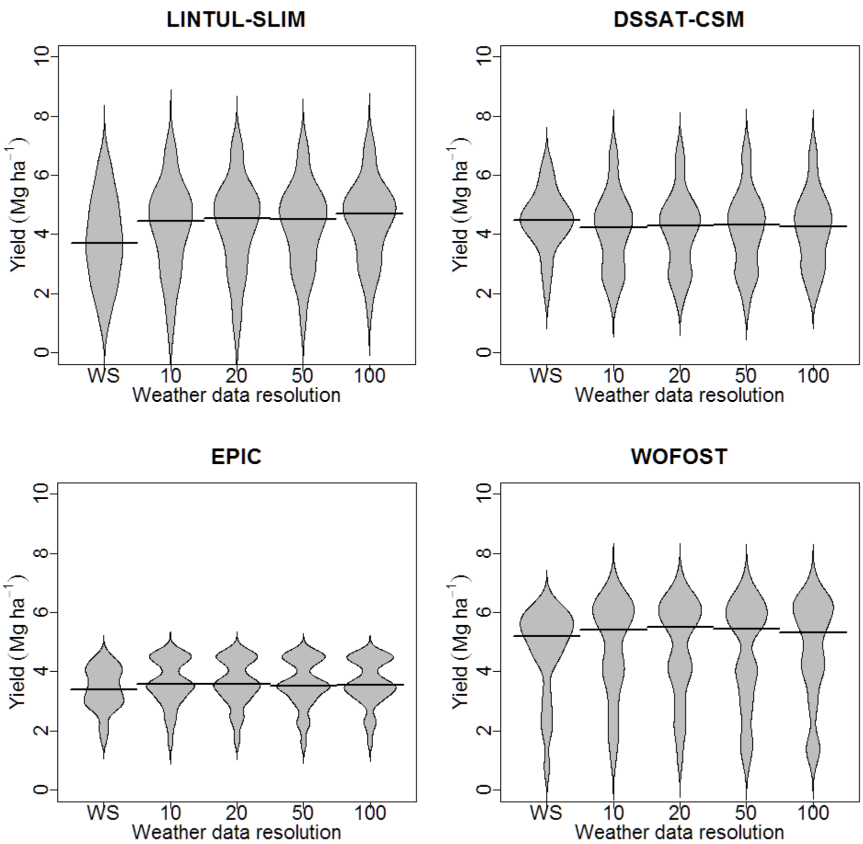

Figure 13. Comparison of frequency distributions of simulated spring barley yields of four crop growth models using 5 weather data resolutions (WS: weather station, 10 = 10 km × 10 km grid cell, 20 = 20 km × 20 km grid cell, 50 = 50 km × 50 km grid cell, 100 = 100 km × 100 km grid cell, horizontal black lines in the bean plot represent the median value of the frequency distribution).

The distributions of yields simulated by DSSAT-CSM showed also a smaller range in the simulated yields, e.g. at 10 km × 10 km resolution, as compared to the other models (Table 6). However, the ranges of simulated yield distributions were in general hardly affected by the resolution of weather data (Figure 13). For example, the range of yields simulated with EPIC was from 2.97 Mg ha^{-1} to 3.29 Mg ha^{-1} (Table 6).

Table 6. Summary statistics of simulated yield (Mg ha^{-1}) for the four models. (Max= highest value; Q= Quartile; 3Q=75% percentile; 2Q=median; 1Q=25% percentile, Min=lowest value; Range=Max-Min).

Model		Weather data resolution				
		Weather Station	Grid cell			
			10 km ×10 km	20 km × 20 km	50 km × 50 km	100 km × 100 km
LINTUL-SLIM	Max	6.83	7.35	7.12	7.05	7.22
	3Q	4.88	5.24	5.30	5.22	5.31
	2Q	3.70	4.44	4.53	4.52	4.70
	1Q	2.63	3.13	3.33	3.28	3.64
	Min	0.82	0.32	0.36	0.85	1.45
	Range	6.01	7.04	6.77	6.20	5.77
DSSAT-CSM						
	Max	6.44	7.02	6.92	7.05	7.04
	3Q	5.01	4.93	4.88	5.00	5.00
	2Q	4.49	4.25	4.30	4.33	4.27
	1Q	3.95	3.08	2.99	3.11	3.20
	Min	2.00	1.72	1.78	1.63	1.99
	Range	4.45	5.31	5.14	5.42	5.04
EPIC						
	Max	4.63	4.76	4.76	4.69	4.64
	3Q	3.99	4.42	4.42	4.41	4.45
	2Q	3.40	3.59	3.59	3.53	3.56
	1Q	2.92	3.11	3.11	3.08	3.19
	Min	1.66	1.48	1.63	1.51	1.58
	Range	2.97	3.29	3.12	3.18	3.06
WOFOST						
	Max	6.23	7.12	7.12	7.11	7.12
	3Q	5.77	6.18	6.22	6.05	6.22
	2Q	5.20	5.41	5.52	5.44	5.31
	1Q	4.43	3.76	3.95	3.34	3.94
	Min	0.76	0.87	0.95	0.88	0.91
	Range	5.47	6.26	6.17	6.24	6.20

Most striking was, however, the difference among the models in the frequency distributions of the simulated yield (Figure 13). Each model exhibited a characteristic distribution or “fingerprint” given by the density trace of the frequency distribution of simulated yields (contour line of the bean plot). When considering distributions of yields simulated with point data (Figure 13), a slight smoothing of the density trace could be noticed. Apart from this, the form of the density trace remained almost unaffected by the extent of weather data aggregation. The yield distributions simulated by LINTUL-SLIM and WOFOST were more spread than those simulated by EPIC and DSSAT-CSM (Figure 13).

3.3.2 Aggregation of inputs *versus* aggregation of outputs

When comparing the density traces of simulated-yield distributions, calculated by using aggregated weather input data at three resolutions, 20 km × 20 km, 50 km × 50 km and 100 km x 100 km (W20, W50 and W100), with each other, only marginal differences could be noticed (Figure 14). Likewise, density traces of distributions of model outputs, i.e. aggregated yields at three resolutions, 20 km × 20 km, 50 km × 50 km and 100 km x 100 km (Y20, Y50 and Y100) showed little differences among themselves (Figure 14). The density trace for Y20, for instance, was smoother and more prolonged towards the lower yielding tail than the resolutions Y50 and Y100. For all resolutions, the aggregation of model outputs (i.e. yields) smoothed the area below the median of the density trace and diminished the range in simulated yields. Nevertheless, the differences between both aggregation strategies were relatively small (Table 7).

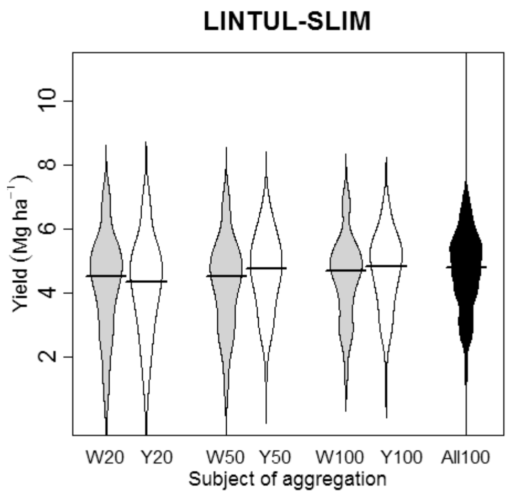

Figure 14. Influence of aggregation of weather inputs (grey fill) and aggregation of simulated yields on yield distributions (white fill) of LINTUL-SLIM. for three resolutions 20 = 20 km × 20 km , 50 = 50 km× 50 km and 100 = 100 km × 100 km . All100 (black fill) = all simulated 10 km x 10 km yields considered.

Table 7. Summary statistics of LINTUL-SLIM yield simulations (Mg ha-1) for different spatial aggregation levels comparing two aggregation strategies aggregation of weather input data (Weather) and aggregation of yields (Yield). The last column contains the summary statistics of non-aggregated yields calculated from 100 grid cells each on a 10 km × 10 km resolution. (Max = highest value; Q= Quartile; 3Q = 75% percentile; 2Q = median; 1Q = 25% percentile, Min = lowest value; Range = Max-Min**).**

Statistics	Aggregation						
	20 km × 20 km grid cell		50 km × 50 km grid cell		100 km ×100 grid cell		100 grid cells (10 km × 10 km)
	Weather	Yield	Weather	Yield	Weather	Yield	No aggregation (all simulated yields)
Max	7.12	7.02	7.05	6.85	7.22	6.85	8.05
3Q	5.30	5.23	5.22	5.48	5.31	5.48	5.58
2Q	4.53	4.34	4.52	4.63	4.70	4.63	4.81
1Q	3.33	3.32	3.28	3.44	3.64	3.44	3.91
Min	0.84	1.13	0.85	1.51	1.45	1.51	1.41
Range	6.29	5.90	6.20	5.35	5.77	5.35	6.64

3.3.3 Distribution of observed yields in the study site

The distribution of individual (not aggregated) observed yields at the study site ranged from 0 to 6.28 Mg ha^{-1} (Table 6). When the observed yields were averaged for each of the four 10 km × 10 km grid cells in which the yield observations were obtained, the density trace became smoother and less spread (see b in Figure 11). When yields were aggregated from the four 10 × 10 km grid cells to a 20 km × 20 km cell (see b in Figure 11), the density trace of the resulting distribution was even smaller and concentrated around the median value 3.7 Mg ha^{-1} (Table 6) showing a bimodal density trace (see c in Figure 11).

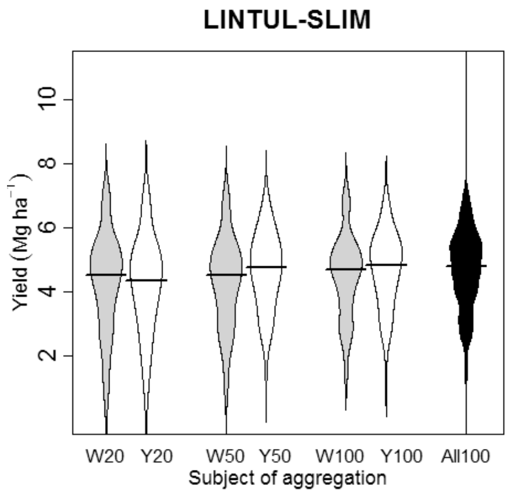

Figure 15. Frequency distributions of observed yields for three levels of spatial aggregation.

Table 8. Summary statistics of observed yields (Mg ha-1). Total sample size: 1204 parcels; minimal sample sizes by grid: 109 fields (Max= highest value; Q= Quartile; 3Q=75% percentile; 2Q=median; 1Q=25% percentile, Min=lowest value; Range=Max-Min).

	No aggregation	Aggregated at 10 km ×10 km	Aggregated at 20 km × 20 km
Max	6.28	4.48	4.28
3Q	4.10	3.94	3.97
2Q	3.50	3.69	3.71
1Q	3.00	3.38	3.43
Min	0.00	2.20	2.56
Range	6.28	2.28	1.72

3.4 Discussion

3.4.1 Choice of weather data resolution

For the selected period and study region, the choice of weather data resolution influenced only marginally the modality and range of the distributions of yields simulated by four crop models differing in detail in the representation of growth processes (Figure 13). The fairly homogeneous topography across the study region leads to a relatively uniform character of weather conditions. It could therefore be argued that, under those conditions, the aggregation of weather input data has only little influence on the simulated yields. Despite the importance of considering the areal effects related to the MAUP when using spatialized weather data, it appears that yield simulations are only little affected by the method of spatialization of weather data in regions where weather conditions are relatively homogeneous. More evident is the difference between aggregated data and single station data. The density traces of the distributions of yields simulated using weather station data appear smoother and more concentrated around the median values than the distributions of yields simulated using weather data from grid cells (Figure 13). It has already been argued that the higher variability displayed by the yields simulated using aggregated weather input data reflects a higher uncertainty introduced by interpolating weather data, for generating weather grid cells (De Wit et al., 2005; Hansen and Jones, 2000; Trnka et al., 2007).

When considering distributions of yields simulated using gridded weather data, only small differences are noticed between the density traces of each resolution step (Figure 13). In former studies (De Wit et al., 2005; Easterling et al., 1998; Olesen et al., 2000; Van Bussel et al., 2011a) the statistical properties of climate data are conserved through a certain resolution range. Comparable to our findings, De Wit *et al*. (2005) found that yields simulated with WOFOST for Germany and France at the national level using aggregated weather at 10 km × 10 km scaled almost linearly with yields simulated at 50 km × 50 km weather data resolution.

The results of the present study are valid primarily for the selected period and study region, and it would be recommendable to perform a similar study in regions where the weather pattern is spatially more heterogeneous.

3.4.2 Fingerprints of models for yield simulations in response to weather

The differences regarding simulated yields among the considered models can be attributed to the way each model processes weather data and calculates weather-related internal crop impact variables such as stresses imposed by temperature and moisture availability (Mearns et al., 1999). In addition, the detail in modelling of light interception and conversion into biomass may explain differences in simulated yield sensitivity to climatic variability (Adam et al., 2011).

Considering both the frequency distribution and summary statistics (such as mean or median values), we get a more comprehensive picture when evaluating regional yield simulations of different models than just looking at a set of few selected statistical indicators. For example, the distributions of yields simulated by LINTUL-SLIM, EPIC and WOFOST using weather station data show lower median values than distributions of yields simulated with grid cell or averaged grid cell data; nevertheless, no major differences can be detected when comparing the corresponding yield distributions. On the contrary, the distribution of yields simulated by DSSAT-CSM with weather station data shows a higher median value than the distributions of yields simulated with grid cell and averaged grid cell weather data. However, when the values of the third quartile (75 percentile) are compared (Table 6), no or few differences (0,13 Mg ha^{-1}) are found between the distributions of yields simulated by DSSAT-CSM.

The selected way of representing distributions in this study – i.e. bean plots –facilitates a better visual assessment of model simulation results. The density trace of the distribution of simulated yields given by the shape of the outer line of the bean plot offers a judgement point for each depicted distribution like a fingerprint for each model (see section 3.3.1) which remains recognizable across different aggregation levels as in this study. The substantial differences in the fingerprints are especially noteworthy among models which are not apparent from summary statistics such as the median.

It would be interesting to understand whether such model specific fingerprints remain recognizable if models are applied across a larger range of environments and whether these fingerprints can be used more systematically in assessing the uncertainty in model simulations. A first attempt in this direction has been made by considering the simulated winter wheat yields from Palosuo *et al*. (2011) for all sites and years of their study and presenting these in the form of bean plots for the models LINTUL-SLIM, DSSAT-CSM and WOFOST also considered in our study (Figure 16). Although the form and extent (density trace) of the distributions of simulated winter wheat yields for the three depicted models do not coincide with their distributions of simulated spring barley yields, there are still similarities which allow visual differentiation and partial identification of each model. The agreement between the bean plots of the two crops is particularly high for DSSAT-CSM which shows almost identical shapes of the corresponding bean plot for spring barley (Figure 13) and winter wheat (Figure 16).

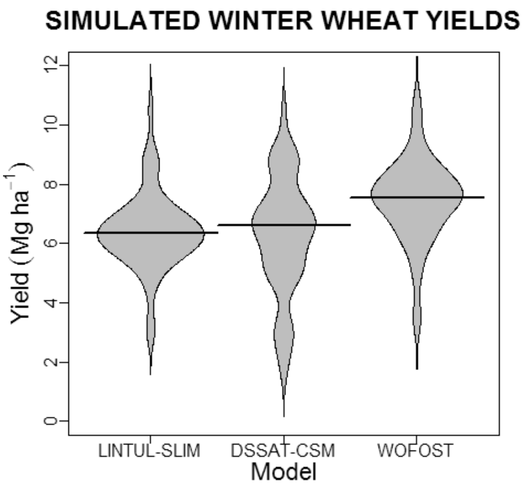

Figure 16. Frequency distribution of simulated winter wheat yields for three models and eight sites (adapted from Palosuo et al., 2011).

Since attempting a thorough causal analysis for the explanation of the form and extent of the model-fingerprints is beyond of the scope of this study, here we provide some main points for discussion, which have already been raised by some previous studies comprising one or more of our four crop models (e.g. Palosuo et al., 2011; Rötter et al., 2012b).

Analysis of frequency distribution for growing period length and biomass (not shown) indicated that growing period length does not seem to be the cause of differences in yield distributions between models. As expected, the shape and partially extent of the bean plots for biomass are almost identical with the yield plots.

The distribution of yields simulated by EPIC is less spread than the distributions of yields simulated with other models. This behaviour could be attributed to the fact that the maximum possible value of LAI (LAI max and LAI min) in EPIC is given as input parameter by the user (Gassman et al., 2004) (see Table 3).

Further analysis also revealed that the way in which water dynamics are calculated seems to represent a main source of differentiation between models based on their frequency distributions of evapotranspiration. Figure 17 depicts the distributions of the values for total evapotranspiration during one growing season (gET) for each model and resolution. Although the outer form (density trace) of the bean plots depicting the gET distributions do not coincide in all the cases with the shape of their corresponding yield distributions (Figure 13), some relationships can be used to characterise yield distributions from the underlying approaches and assumptions to calculate water dynamics. All models simulated water dynamics utilizing the capacity approach, but the level of detail describing the soil profile differs considerably among models. For instance, WOFOST (Boogaard et al., 1998; Van Diepen et al., 1989) describes soil as a two homogeneous layer profile, whereas in EPIC (Izaurralde et al., 2006) the soil profile is represented by up to 10 layers. Interestingly, the simpler approach in WOFOST causes less variability in gET as compared to the more complex, multi layer soil modules in EPIC and DSSAT-CSM. Also, the variability in gET in WOFOST seems to be less important for the characteristic yield distribution depicted for the model as compared to the other models for which the bean plots for gET and yield are more similar. Whether in WOFOST the variability in gET is less important for determining the variability in yields as for the other models needs further evaluation. However, the present form of presenting results can give hints on further analysis required.

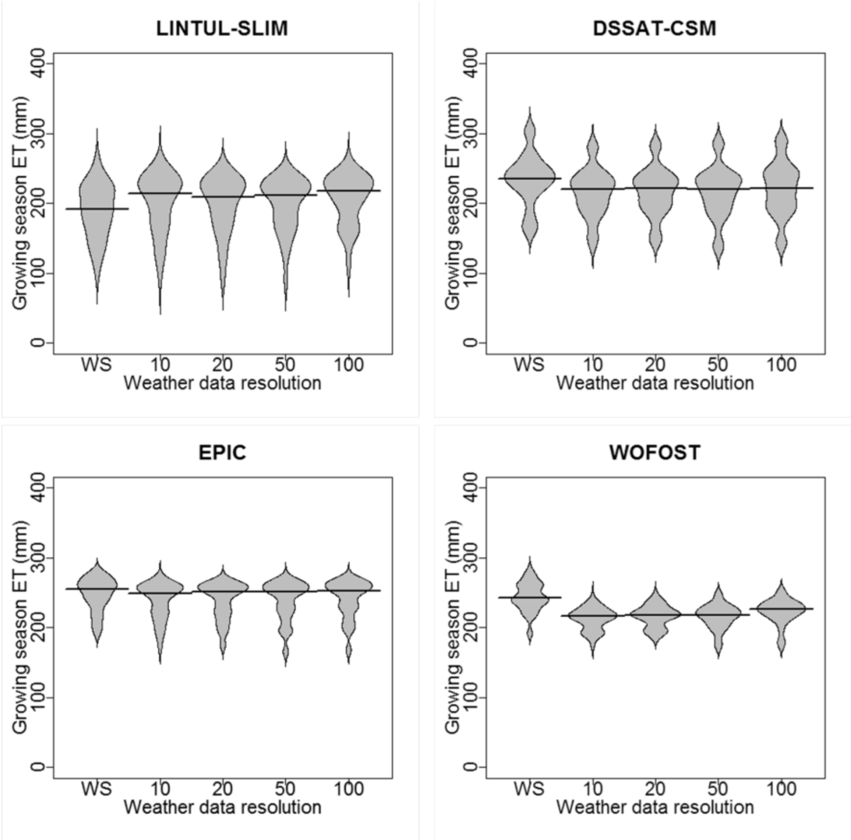

Figure 17. Comparison of frequency distributions of simulated total growing season evapotranspiration of four crop growth models using 5 weather data resolutions (WS: weather station, 10 = 10 km × 10 km grid cell, 20 = 20 km × 20 km grid cell, 50 = 50 km × 50 km grid cell, 100 = 100 km × 100 km grid cell, horizontal black lines in the bean plot represent the median value of the frequency distributions.

3.4.3 Aggregation issues

Weather

The strategy of aggregating weather data chosen in this study refers to averaging of data across a spatial unit. Former studies (Easterling et al., 1998; Hansen and Jones, 2000) concluded that averaging weather variables such as temperature and precipitation over space, might influence negatively their daily variability. Since the main objective of this study was to test the influence of spatial weather data aggregation, no segregation of individual variables was undertaken. Nevertheless, precipitation seems to play a decisive role as yield influencing factor across aggregation levels. Hansen and Jones (2000), among others, found that by averaging weather data, simulated yields might be overestimated. On the one hand, more frequent but less intense precipitation events might not recharge soil water reserves in deeper layers and favour augmented evaporation. On the other hand, more frequent precipitation events could reduce the duration of dry spells between rain events and decrease the probability of water stress. Thus, simulated soil water balance components and their specific relations to simulated yields might be positively influenced by weather data aggregation. Results of the present study suggest that the effect of weather data aggregation is relatively small and depends on the model.

Simulated yields

The lower value range for the distributions of aggregated yields (Table 8) might be interpreted as a result of variability loss, caused by aggregation (De Wit et al., 2005). Again, depicting the yield distributions considering density traces facilitates the evaluation process. Using this type of visualization shows that the model specific fingerprints also remained when output data, i.e. simulated yields, were aggregated for the selected resolutions (Figure 13). However, to which extent the small differences between the two aggregation strategies (i.e. aggregating model input data *versus* aggregating model output data) apply to other regions where weather data are more spatially heterogeneous awaits further testing.

Observed yields

The shape of yield distributions (density trace) for the selected period and site changed when observed yields were aggregated at 10 km × 10 km and 20 km × 20 km resolution. The present study did not focus on methods for aggregating observed yields. Other methods different from the weighted-average might be more adequate for aggregating observed yields (Hansen and Jones, 2000). However, according to our results, the effect of spatial aggregation has to be considered when utilizing observed yields especially for the calibration process, since uncertainties in model results are not only related to model deficiencies but also to error introduced through insufficient or misleading calibration (Palosuo et al., 2011).

3.5 Conclusion

In the selected study region and period (12 years), spatial aggregation of weather input data influenced only marginally the shape and extent of simulated yield distributions visualised in the form of bean plots of four crop models. Differences in yield distributions were most striking between models rather than between aggregation levels. We therefore propose that crop models can be typified and further evaluated according to their specific yield distribution form and range - a so called fingerprint - which determines the probability path of simulated yields under a range of weather conditions in a region. This can be extended to the underlying processes to better inform about relationships between the variability of processes and yield. However, further evaluation will be required to understand the robustness of a model’s fingerprint across a larger range of conditions including other factors such as soil and management, and the relationships to the underlying processes in order to better explain difference in fingerprints among models. Nevertheless, we believe that it is more advantageous to evaluate model performance considering also the frequency distributions than relying on selected summary statistics such as mean, median or standard deviation only.

Our results also support recent findings that a multi-model use should be the preferred option when assessing climate impacts on regional crop yields.

We finally would like to stress the need of careful evaluation when aggregating simulated yields as compared to aggregating model input data as yield distributions can change depending on the aggregation level. Finally we recommend using observed site-specific yields to better understand the yield distribution within an area as this can be modified through aggregation.

Chapter 4

‘Fingerprints’ of four crop models as affected by soil input data aggregation

The present chapter has been published as:

Angulo, C., Gaiser, T., Rötter, R.P., Børgesen, C.D., Hlavinka, P., Trnka, M., Ewert, F., 2014. “Fingerprints” of four crop models as affected by soil input data aggregation. Eur. J. Agron,. 6: 35–48. doi:10.1016/j.eja.2014.07.005

4. 'Fingerprints' of four crop models as affected by soil input data aggregation.

4.1 Introduction

The regional application of plot-scale mechanistic crop growth simulation models (further on referred to as crop models) is regarded critical, mainly because of the scale change issues inherent to the model spatialisation process (Ewert et al., 2011b; Hansen and Jones, 2000) and its requirements for a high amount and quality of input data (Faivre et al., 2004; Leenhardt et al., 2006). Nevertheless, crop models have become a standard tool to assess plant production and crop productivity due to their explanatory character and low cost applicability also for large area applications (e.g. Angulo et al., 2013a; Batchelor et al., 2002; Ewert et al., 2011b; Reyenga et al., 1999; Rötter et al., 2011a; Rötter et al., 2011b; Tubiello et al., 2007; Van Ittersum et al., 2008; Wassenaar et al., 1999). Crop models typically require input data related to weather, soil characteristics and crop management (Adam et al., 2012). Since the heterogeneous spatial distribution of soil properties is an important source of yield variability (e.g. Batchelor et al., 2002; Mignolet et al., 2004; Wassenaar et al., 1999), it is crucial to have sound soil input data available in order to obtain plausible regional yield simulations.

There are conflicting results concerning the influence of the spatial resolution of soil input data on simulated yields. On the one hand, it has been found that for relatively small areas like the Hérault-Libron-Orb valleys in France (approx. 1 200 km^2), the variability of winter wheat yields simulated by EuroACCESS is strongly affected by the soil input data variability (Wassenaar et al., 1999). The importance of soil input data variability is attributed by the authors to the size of the region which favours the importance of soil data variability over a less variable climate and conditions of water limitation as typical for Mediterranean regions. In a national yield assessment study in Denmark, the capability of the crop model CLIMCROP to reproduce the spatial winter wheat yield variability was reduced when soil data at low resolution (dominant soil in a county) were used as input as compared to the highest resolution (1: 50 000 soil map) (Olesen et al., 2000). On the other hand, the consideration of different soil input data resolutions: field-measured soil data and soil data bases with a mapping scale of 1:20 000 and 1:250 000, did not play an important role as source of uncertainty when simulating non irrigated sorghum yields with EPIC in the Great Plains (Niu et al., 2009). The authors suggest that the low introduction of uncertainty by soil input data might be related to the fact that the properties of the dominant soils were identical for field measured data and the 1:250 000 soil map and very similar to the 1:20 000 soil map. In a different study Folberth *et al.* (2012) concluded that the model GEPIC is less sensitive to the resolution of soil input data in comparison to the resolution of management (irrigation) and climate input data for simulating grain maize yields in the United States, confirming earlier results by Easterling et al. (1998).

The mentioned studies focused mainly on the influence of input data resolution on the predictive power of crop models and have not explicitly evaluated the influence of different levels of spatial resolution of soil input data on simulated yield distributions and the correspondent underlying causes. Moreover, all of those studies have only considered a single crop model approach. However, according to the results of chapter 2 (Angulo *et al.* 2013b) differences among models were more pronounced than those among scaling methods. This is in line with multi-model field scale studies (e.g. Palosuo et al., 2011; Rötter et al., 2013b; Rötter et al., 2012b) and suggests that adopting a multi-model approach in regional yield assessment studies might allow to quantify the uncertainty in (specially future) simulated yields introduced by the crop growth simulation approach (Asseng et al., 2013). The study by Angulo et al. (2013b) also points to the usefulness of considering yield distributions, so called ‘model fingerprints’, to characterize and evaluate the frequency distributions of simulation results of crop models. Following up on this research our study aims to answer two main research questions: 1) what is the influence of the spatial resolution of soil input data on the distributions of simulated yields? 2) What are the differences in the behaviour of crop models, which differ in model approach and detail to different spatial resolutions of soil input data for simulating regional (i.e. county level) yields? To answer these questions we considered four crop models (SIMPLACE<LINTUL-SLIM>, DSSAT-CSM, EPIC, DAISY) applied in two regions in Germany with different climate and soil characteristics. The importance of soil input data aggregation was explored for different resolutions. In extension of the study by Angulo *et al.* (2013b), we analysed crop models behaviour for simulated yield and total growing season evapotranspiration considering frequency distributions ‘fingerprints’ to get more elaborated insights into the uncertainty of model simulations.

4.2 Materials and Methods

4.2.1 Crop models

Four crop models were used in this study: The SIMPLACE<LINTUL-SLIM> solution of the modelling platform SIMPLACE (Scientific Impact Assessment and Modelling Platform for Advanced Crop and Ecosystem Management) (Addiscott and Whitmore, 1991b; Angulo et al., 2013b; Gaiser et al., 2013), DSSAT-CSM (Jones et al., 2003a), EPIC (Gassman et al., 2004) and DAISY (Hansen et al., 2012). All models have already been calibrated and applied to simulate yields of winter and spring cereals for German conditions (Angulo et al., 2013a; Gaiser et al., 2009; Gaiser et al., 2013; Palosuo et al., 2011; Rötter et al., 2012b). Since our study did not focus on analysing the predictive power of the models but rather on their behavioural response to different soil data resolutions, no calibration in a strict sense was undertaken but only a partial validation in order to verify the plausibility of the models’ results. For this purpose the four models were applied to simulate winter wheat yields using phenological and management data for the years 2003 and

2004 from the German variety trials at the station Kerpen-Buir located in North Rhine-Westphalia (Figure 18) (LSV, 2012). The differences between simulated and observed yields ranged from 1.55 Mg ha^{-1} for EPIC (in 2003) to 0.11 Mg ha^{-1} for SIMPLACE<LINTUL-SLIM> (in 2004). These values correspond to a relative RMSE of 2% for SIMPLACE<LINTUL-SLIM>, 6% for EPIC, 11% for DSSAT-CSM and 9% for DAISY of the simulated yields in relation to the observed yields. In a blind test with very restricted calibration, Palosuo *et al.* (2011) obtained higher deviations between simulated and observed winter wheat yields and concluded that despite a total variation of 18% the multi-model mean estimates could still reproduce observed yields satisfactorily.

Table 9. Major processes determining crop growth and development of the models applied in this study (Modified from Palosuo et al., 2011).

Model	SIMPLACE <LINTUL-SLIM>	DSSAT-CSM	EPIC	DAISY
Version	220	4.6.0.8	0509	4.01
Leaf area development and light interception [a]	D	D	S	D
Light Utilization [b]	RUE	RUE	RUE	P-R
Yield formation [c]	Y(Prt)	Y((GnGw,En.Prt)B)	Y(HI,B)	Y(B,Prt)
Crop Phenology [d]	f(T,DL,V)	f(T,DL,V)	f(T,DL)	f(T,DL,V)
Stresses involved [e]	W	W,N	W,N,P	W,N
Soil water dynamics [f]	C	C	C	R
Evapotranspiration [g]	PM	PT	PM	MK

[a] Leaf area development and light interception: S = simple or D = detailed approach.

[b] Light utilization or biomass growth: RUE = Simple (descriptive) Radiation use efficiency approach, P-R = Detailed (explanatory) Gross photosynthesis-respiration.

[c] Y(x) Yield formation depending on: HI = fixed harvest index, B = total (above ground) biomass, Gn = number of grains, Prt = partitioning during reproductive stages, Gw=grain weight, En=earn number.

[d] Crop phenology is a function (f) of: T = temperature, DL = photoperiod (day length), V = vernalisation; O = other water/nutrient stresses effects considered.

[e] Stresses involved: W = water, N = nitrogen stress, P = phosphorus stress (Not considered in this paper).

[f] Water dynamics approach: C = capacity approach, R = Richards approach.

[g] Method to calculate evapo-transpiration: P = Penman, PM = Penman-Monteith, PT = Priestley-Tailor, MK= Makkink.

The main characteristics of the four models used in the present study are summarized in Table 9. According to the classification suggested by Adam *et al.* (2011), considering light utilization and light interception as main differentiators between models, three types of models can be distinguished in our study: EPIC is a fairly simple crop model since it adopts the concept of radiation use efficiency (RUE) which is a summarized light utilization approach (Monteith and

Moss, 1977) and a summarized leaf area index (LAI) dynamics approach based on a forcing function (Williams et al., 1983). SIMPLACE<LINTUL-SLIM> and DSSAT-CSM can be considered crop models of intermediate complexity since they use the summarized RUE light utilization approach and a detailed LAI calculation approach driven by the development stage of the crop and development specific partitioning fractions (Spitters, 1990). Finally, DAISY can be considered a fairly detailed crop model which applies a detailed light utilization approach calculating gross photosynthesis and maintenance and growth respiration (Van Ittersum et al., 2003a), and a detailed LAI calculation approach.

All models have modules to calculate soil water dynamics. DAISY uses the detailed Richards approach for soil water movement, the other three use a simpler capacity or tipping bucket approach (Van Ittersum et al., 2003a). To calculate potential evapotranspiration SIMPLACE<LINTUL-SLIM> and EPIC apply the Penman-Monteith equation according to Allen *et al,* (1998a) and Monteith and Greenwood (1986) respectively; and in this study, in DSSAT-CSM the Priestly-Taylor equation (Priestley and Taylor, 1972) is applied, while in DAISY it is the Makkink equation (Makkink, 1957).

4.2.2 Study areas

For the present study we considered seven counties in the Federal State of North-Rhine Westphalia (Figure 18) located in two contrasting regions in terms of elevation, geomorphology and climate (Table 3) according to the German agricultural and forest zonation system BKR (“Boden-Klima-Raum”, German for “Soil-Climate-Zone”) (Rossberg et al., 2007). The counties Aachen, Düren, Erftkreis and Heinsberg are located in the BKR 141 known as “Julicher Börde”. This region spreads over large plain formed by the Rhine river and marine influence in the Paleozoic period and is characterized by mild temperatures. The counties Märkischer Kreis, Olpe and Siegen-Wittgenstein are located in the region “Sauerland” which is a mountainous and cooler area (see Table 3).

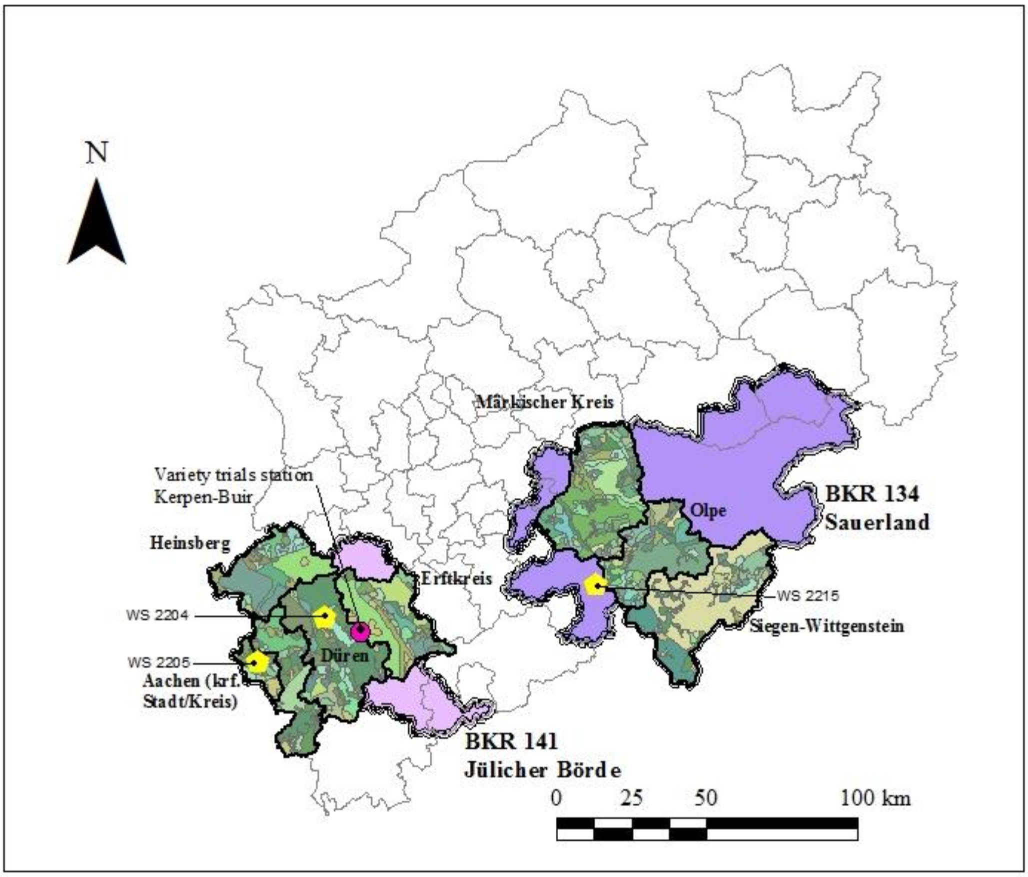

Figure 18. Location of the selected seven counties, variety trials station Kerpen-Buir, and correspondent three weather stations in the German Federal State of North-Rhine-Westphalia (WS=Weather station, BKR=Soil-climate-zone).

4.2.3 Soil Data

The geomorphologic setting of the two regions differs in terms of relief and petrography. The Jülicher Börde region is a large plain covered by Aeolian loess deposits with a depth of several meters. A few remnant outcrops of the tertiary occur within the plain but they are not used as cropland. In contrast to the Jülicher Börde region, the Sauerland region is a mountainous area with undulated topography. The soil parent materials are consolidated sedimentary rocks (shale, mudstones) of Paleozoic origin which are occasionally covered by shallow Loess deposits with variable depth, which are the preferred cropland soils.

Spatial resolution of soil data

In practice, soil information at high resolution (mapping scale higher than 1:200000) is scarce due to the extremely high cost of soil mapping. Our study envisaged to clarify if there is a minimum soil data resolution which is necessary to adequately reproduce crop yields in the context of regional crop modelling applications. The commonly used technique to generate soil maps (i.e. generalization in cartographic terms) from higher to lower resolution, is to unite (larger) areas that share similar geological and geo-morphological characteristics (soil sub-units) into a generalized class characterized by a dominant soil unit represented by a typical soil profile (Leenhardt et al., 1994). Accordingly, the following three spatial resolutions of input soil data were tested for the present study.

The **first** soil data resolution (**res1**) is based on the most detailed soil map with the highest resolution at a scale of 1:50 000 provided by the Geological Service of North-Rhine Westphalia (BK50, 2004). The map contains information about the distribution of approximately 7000 soil units for the whole State of North-Rhine Westphalia. Each soil unit is characterized by a representative soil profile description which reports the following properties: number and depth of soil layers, texture and gravel content of each layer, depth of ground water table, soil type and sub-type according to the German Soil Classification. In **res1** the spatial distribution of each mapping unit (at the soil sub-type level) with its respective soil properties in the two regions was considered. The **second** soil data resolution (**res2**) was obtained by aggregating the mapping units of the soil sub-types to the soil type level, which corresponds approximately to a mapping scale of 1:300 000. For the aggregation procedure we took into consideration only the soil sub-type unit with the highest spatial coverage within each of the seven counties and we assumed that this dominant soil sub-type unit with its specific soil profile is representative for all other sub-types belonging to the same soil type. The **third** soil data resolution (**res3**) is based on the German soil map at the reconnaissance level at a mapping scale of 1:1 000 000 (Hartwich et al., 1995). The dominant soil types in each county where extracted by overlaying the reconnaissance map with the boundaries of the seven counties. In order to keep consistency in soil properties for the same soil type in the three spatial resolutions, we did not use the soil profile descriptions for the different soil types as suggested by Hartwich *et al.* (1995). For each soil type on the reconnaissance map we rather used the soil profile description of the corresponding dominant soil sub-type on the most detailed soil map of scale 1:50 000 (BK50, 2004). Table 10 presents an overview of the number of soil profiles used for each resolution. For all three resolutions only the soil profiles occurring on cropped land were taken into consideration. The water holding capacity of the soils, required as soil input data for three models, was estimated based on the texture class information applying tabular pedotransfer functions developed for German soils (AG-Boden, 2005) (Appendix 1). Additionally, for the model DAISY, pedotransfer functions based on the data base HYPRES (Wösten et al., 1999) were applied to calculate the required van Genuchten/Mualem parameters (θ_s, K_s, α, l and n).

Table 10. Number of soil profiles per county and resolution with respective weather station considered for yield simulations.

County name	Number of profiles according to resolution:			Weather station ID (code)
	res1 1:50.000 Soil sub-unit	res2 1:300.000 Soil unit	res3 1:1.000.000 Dominant soil unit	
Aachen	88	13	6	2205
Dürren	241	23	7	2204
Erftkreis	132	17	4	2205
Heinsberg	217	27	5	2204
Sum four counties in Jülicher Börde	*678*	*80*	*22*	
Märkischer Kreis	122	17	6	2215
Olpe	90	16	5	2215
Siegen-Wittgenstein	90	13	3	2215
Sum three counties in Sauerland	*302*	*46*	*14*	

4.2.4 Weather data

Daily weather data including global solar radiation (MJ m^{-2} d^{-1}), maximum air temperature (°C), minimum air temperature (°C), precipitation amount (mm d^{-1}) and wind speed (m s^{-1}) for the period from 1994 to 2008 provided by the German weather service (DWD, 2010) were used. Two weather stations were considered in the Jülicher Börde according to their proximity to the respective counties: the station Aachen WEWA (Code-2205) for the counties Aachen and Erftkreis; and the station Jülich Kernforschungsanlage (Code-2204) for the counties Düren and Heinsberg. In the Sauerland, the station Reichshof-Eckenhagen (Code-2215) was used for all counties (Olpe, Siegen-Wittgenstein and Märkischer Kreis) (Figure 18). Station characteristics are given in Table 11. Precipitation for the period from 1994 to 2008 in both regions is abundant enough over the growing period (>450mm) excepting for the year 1996.

Table 11. Weather statistics of the three weather stations used in this study for the period 1995-2008.

	Jülicher Börde		*Sauerland*
Station name	Jülich Kenrforschungsanlage	Aachen WEWA	Reichshof-Eckenhagen
Station ID (code)	2204	2205	2215
Latitude	50.91	50.78	51.00
Longitude	6.41	6.10	7.70
Altitude (m.a.s.l.)	91	202	350
Mean annual precipitation (mm)*	705	812	1183
Mean annual maximum temperature (°C)	14.95	14.58	12.83
Mean annual minimum temperature (°C)	6.58	7.23	5.72
Mean growing season precipitation (mm)+	578	664	970
Mean growing season maximum temperature (°C)	13.95	13.55	11.79
Mean growing season minimum temperature (°C)	5.8	6.34	4.8

*Value represents not only rain but also snow.

+ Period from 1994 to 2008.

4.2.5 Yield simulations

For each model three yield simulation steps corresponding to the three spatial resolutions of soil input data were performed (Table 10). In each step yields were simulated for 14 years from 1995 to 2008. All models were configured to use the optimum nitrogen fertilizer amount for the study region and the 15th of October as yearly sowing date.

The distributions of both simulated yields and simulated total growing season evapotranspiration for each model were assessed in form of bean plots. Bean plots similarly to box and whisker plots depict the range of a data sample without further assumptions of the distribution or median. Additionally the contour line of the bean plot represents a density trace which offers an insight of the frequency distribution of the sample. For the calculation of the density trace the normal (Gaussian) kernel was used. Kampstra (2008) offers a detailed description of the implementation of bean plots.

In order to quantitatively compare the relative importance the of two sources of uncertainty: model choice and resolution of input data, we calculated for each model the coefficient of variation (CV) of simulated results and the coefficient of variation of the root mean square error (CV(RMSE)) of the yields simulated with res2 and res3 compared to res1 (see 4.4.3). The CV is calculated by dividing the standard deviation of the simulated yields by their mean. Equation (3) in the section 2.2.5.1 describes the calculation of RMSE. The CV(RMSE) is the result of dividing the RMSE of yields with their mean. The yields simulated with vres1 were used as observed values of the calculation of RMSE.

4.3 Results

4.3.1 Influence of soil data resolution on simulated yields

For all models and counties considered in this study only a minimal influence of the spatial resolution of soil input data on the shape of the density traces of simulated yields was found. Since the distribution of simulated yields between the counties located in one region were very similar, results are not presented for individual counties but are summarized for the two regions as described in section 2.2. No aggregation of results was undertaken but simulated yields for all soil units within the four counties in the Jülicher Börde and three counties in the Sauerland were used to calculate the corresponding bean plots (Figure 19).

The choice of spatial soil data resolution affected minimally the extent of the simulated yield distributions. When using a coarser resolution (res3) the simulated yield range was smaller for all counties and models compared to the range using the highest resolution (res1) (Table 12 and Table 13). Only in the specific case of the yields simulated by DAISY for the county Siegen-Wittgenstein the range remained equal for all resolutions (Table 13). When considering all models and counties, there is no recognizable difference in the median values with respect to the spatial resolution of soil input data.

4.3.2 Interrelation between model and aggregation

There is a remarkable difference of the shapes and extents of the simulated yield distributions between the four models and between the regions (Figure 19). For instance, in the region Jülicher Börde (BKR 141) the ranges of the simulated yield distributions are res1: 7.1, res2: 6.3 and res3: 4.4 Mg ha^{-1} for SIMPLACE<LINTUL-SLIM>, while for EPIC the ranges of the respective distribution are: res1: 9.6, res2: 9.1 and res3: 5.14 Mg ha^{-1}. Also, judging from the form of the bean plots, which represents the probability density trace of the simulated yields, the most probable value for yields simulated by SIMPLACE<LINTUL-SLIM> is around 7.4 Mg ha^{-1} while for the yields simulated by EPIC this is over 8.0 Mg ha^{-1}.

Similarly, when considering the distributions of simulated total growing season evapotranspiration, the differences between models in terms of range and shape of the distributions are evident (Figure 20). When comparing both simulated yield and simulated total growing season evapotranspiration for each model, no systematic correspondence between the shapes or extent of both outputs for the same model can be found. On the one hand, the ranges and mean values of both the distributions of yields and the distributions of total growing season evapotranspiration for the two regions simulated by EPIC are for each resolution the largest of all models (Table 12 and Table 13). On the other hand, although the range of yields simulated by DSSAT-CSM is the second highest of all models, the corresponding range values of simulated total growing season evapotranspiration is the lowest of all.

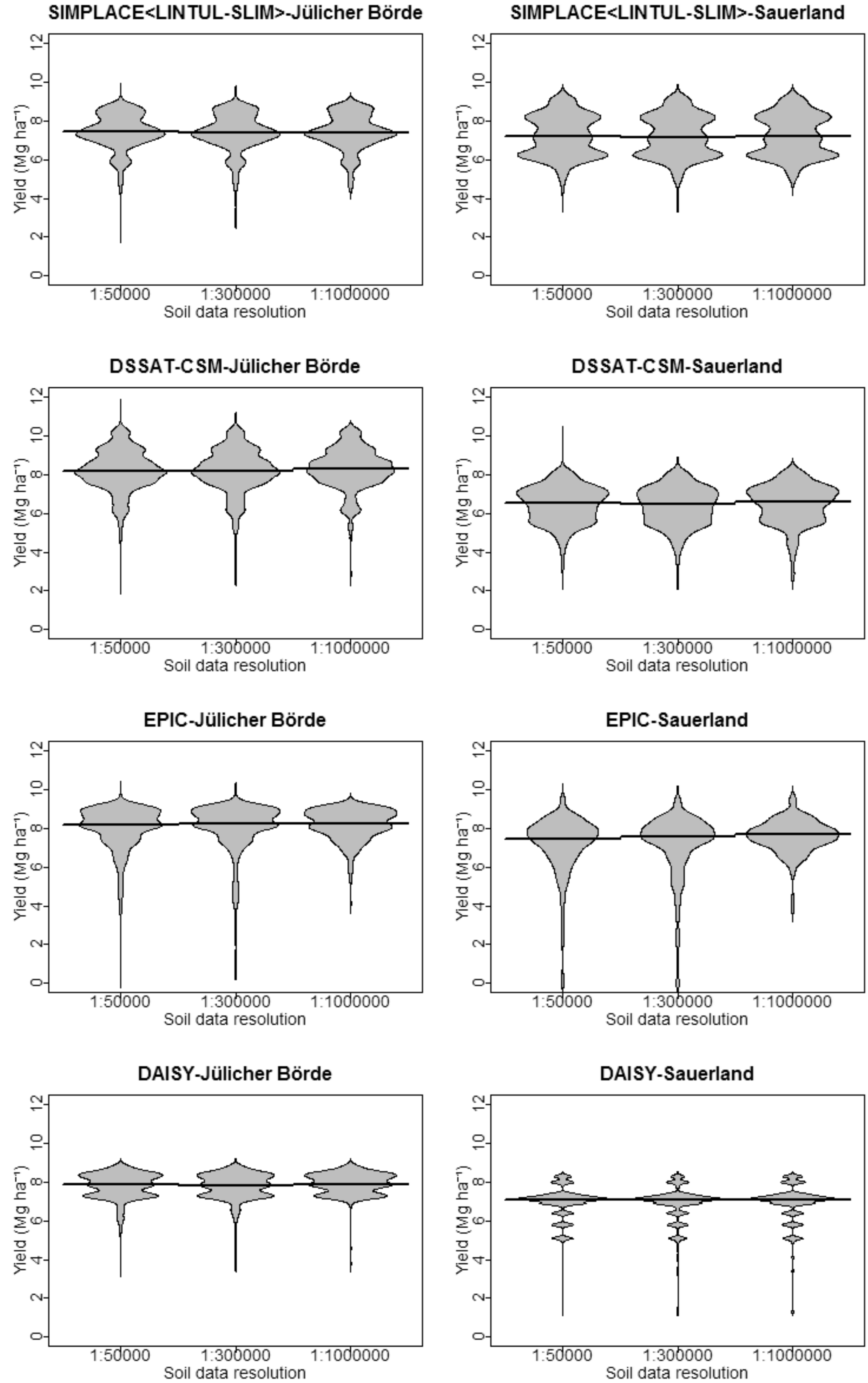

Figure 19. Comparison of frequency distributions of simulated yields of four crop growth models using three soil data resolutions (horizontal black lines in the bean plot represent the median value of the frequency distribution).

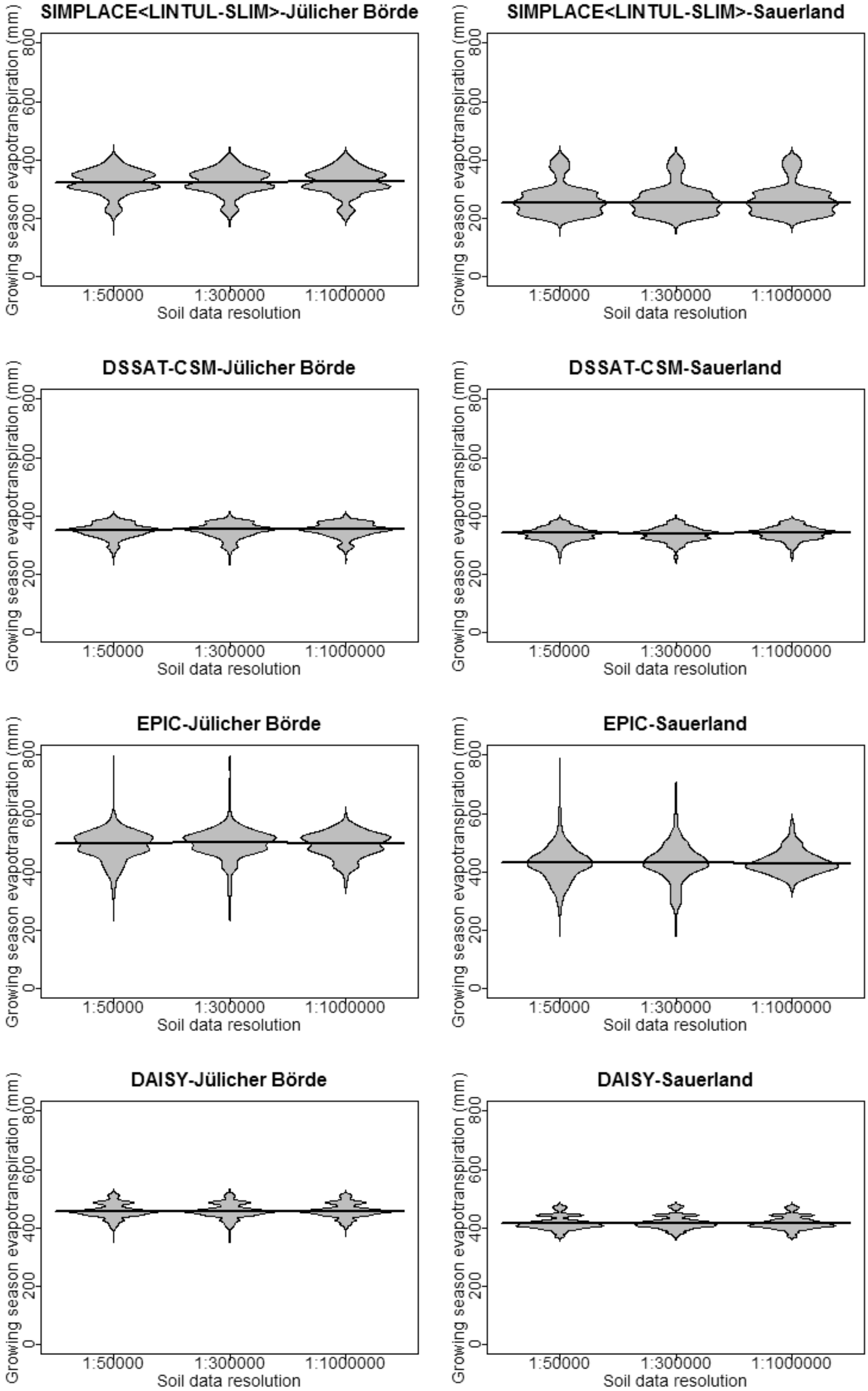

Figure 20. Comparison of frequency distributions of simulated total growing season evapotranspiration of four crop growth models using three soil data resolutions (horizontal black lines in the bean plot represent the median value of the frequency distributions).

Table 12. Summary statistics of the distributions of winter wheat yields in four counties in the Jülicher Börde (BKR 141) simulated by four crop models using soil input data at three spatial resolutions : 1.- 1:50.000, 2.-1:3.000.000; 3.-1:1.000.000 (Max= highest value; Q= Quartile; 3Q=75% percentile; 2Q=median; 1Q=25% percentile, Min=lowest value; Range=Max-Min).

County	Aachen			Düren			Erftkreis			Heinsberg			Jülicher Börde KR141 (summary all 4 counties)		
Resolution	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
SIMPLACE<LINTUL-SLIM>															
Max	9.0	8.9	8.9	9.4	9.3	8.7	8.9	8.9	8.9	9.4	9.3	8.6	9.38	9.25	8.92
3Q	8.3	8.3	8.3	8.0	8.0	8.1	8.3	8.3	8.4	8.1	8.0	7.8	8.14	8.1	8.11
2Q	7.3	7.3	7.3	7.5	7.5	7.5	7.3	7.3	7.4	7.5	7.5	7.4	7.45	7.4	7.39
1Q	6.9	6.9	6.9	7.0	7.0	7.0	6.9	6.8	6.9	7.1	7.0	7.0	7.01	6.96	6.93
Min	2.9	4.5	4.5	2.3	4.5	4.6	3.0	3.0	4.9	3.9	4.3	5.3	2.29	2.98	4.54
Range	6.2	4.4	4.4	7.1	4.7	4.1	6.0	5.9	4.0	5.5	5.0	3.3	7.09	6.27	4.38
DSSAT-CSM															
Max	10.2	10.2	10.2	10.5	10.4	10.2	11.3	10.2	10.2	10.7	10.6	10.2	11.32	10.62	10.23
3Q	9.0	8.9	8.9	8.6	8.6	8.6	9.1	9.2	9.2	8.6	8.6	8.6	8.85	8.85	8.86
2Q	8.5	8.6	8.6	8.1	8.1	8.2	8.6	8.7	8.7	8.1	8.0	8.1	8.2	8.21	8.33
1Q	7.9	7.9	7.9	7.6	7.6	7.7	7.8	7.9	7.8	7.6	7.5	7.6	7.64	7.64	7.65
Min	2.8	5.6	6.0	3.1	3.5	6.1	2.4	5.2	5.3	2.9	2.9	2.9	2.44	2.88	2.88
Range	7.4	4.6	4.1	7.4	6.9	4.1	8.9	5.0	4.9	7.8	7.7	7.4	8.88	7.74	7.35
EPIC															
Max	9.6	9.4	9.3	9.9	9.8	9.2	9.9	9.4	9.3	9.9	9.6	9.2	9.88	9.82	9.29
3Q	8.7	8.8	8.8	8.7	8.8	8.6	8.8	8.8	8.8	8.9	8.9	8.9	8.76	8.83	8.79
2Q	8.0	8.2	8.5	8.1	8.2	8.0	8.2	8.5	8.5	8.3	8.3	8.4	8.21	8.26	8.24
1Q	7.2	7.6	7.9	7.3	7.4	7.3	7.6	7.8	7.9	7.9	7.8	8.1	7.5	7.7	7.79
Min	0.3	4.2	4.2	3.6	3.7	5.2	1.6	4.6	6.1	0.7	0.7	6.3	0.3	0.71	4.15
Range	9.3	5.2	5.1	6.3	6.1	4.0	8.3	4.8	3.2	9.2	8.9	2.8	9.58	9.11	5.14
DAISY															
Max	8.5	8.5	8.5	8.8	8.8	8.8	8.5	8.5	8.5	8.8	8.8	8.8	8.8	8.8	8.8
3Q	8.2	8	8.2	8.4	8.3	8.4	8.2	8.2	8.35	8.3	8.3	8.3	8.3	8.3	8.4
2Q	7.8	7.8	7.8	7.9	7.9	8.2	7.8	7.8	7.9	7.9	7.9	8.1	7.9	7.8	7.9
1Q	7.3	7.3	7.5	7.3	7.3	7.4	7.4	7.5	7.7	7.3	7.3	7.4	7.3	7.3	7.5
Min	3.8	3.8	3.8	4.4	5.4	7.2	3.5	5.7	7.3	4.0	4.0	7.1	3.5	3.8	3.8
Range	4.7	4.7	4.7	4.4	3.4	1.6	5.0	2.8	1.2	4.8	4.8	1.7	5.3	5	5

Table 13. Summary statistics of the distributions of winter wheat yields of three counties in Sauerland (BKR 134) simulated by four crop models using soil input data at three spatial resolutions: 1.- 1:50.000, 2.-1:3.000.000; 3.-1:1.000.000 (Max= highest value; Q= Quartile; 3Q=75% percentile; 2Q=median; 1Q=25% percentile, Min=lowest value; Range=Max-Min).

County	Märkisher Kreis			Olpe			Siegen-Wittgenstein			Sauerland (BKR 134) (summary all 3 counties)		
Resolution	1	2	3	1	2	3	1	2	3	1	2	3
SIMPLACE<LINTUL-SLIM>												
Max	9.1	9.1	9.1	9.1	9.1	9.1	9.1	9.1	9.1	9.12	9.12	9.1
3Q	8.2	8.2	8.2	8.1	8.1	8.2	8.1	8.1	8.0	8.16	8.13	8.17
2Q	7.2	7.2	7.3	7.2	7.1	7.0	7.2	7.1	7.0	7.17	7.14	7.19
1Q	6.3	6.3	6.3	6.3	6.3	6.3	6.3	6.3	6.3	6.29	6.29	6.29
Min	4.1	4.1	5.3	4.1	4.1	4.9	4.2	4.2	5.2	4.1	4.1	4.92
Range	5.0	5.0	3.8	5.0	5.0	4.2	4.9	4.9	3.8	5.02	5.02	4.18
DSSAT-CSM												
Max	8.1	8.1	8.1	8.2	8.2	8.1	9.8	8.1	8.1	9.75	8.21	8.11
3Q	7.1	7.1	7.3	7.1	7.1	7.2	7.2	7.1	7.2	7.15	7.09	7.26
2Q	6.5	6.5	6.8	6.5	6.4	6.5	6.6	6.5	6.5	6.51	6.47	6.59
1Q	5.7	5.7	6.1	5.7	5.5	5.5	5.9	5.7	6.1	5.77	5.54	5.82
Min	2.9	3.6	5.1	2.8	2.8	2.8	3.5	3.9	5.1	2.79	2.79	2.79
Range	5.2	4.5	3.1	5.4	5.4	5.3	6.3	4.2	3.0	6.96	5.42	5.32
EPIC												
Max	9.6	9.6	9.6	9.7	9.5	9.5	9.7	9.5	8.7	9.66	9.55	9.55
3Q	8.1	8.2	8.2	8.1	8.1	8.1	8.0	8.0	7.8	8.04	8.13	8.14
2Q	7.5	7.7	7.9	7.6	7.5	7.6	7.4	7.5	7.3	7.47	7.55	7.66
1Q	6.7	7.1	7.4	6.7	6.3	7.1	6.5	6.4	6.9	6.63	6.65	7.2
Min	0.0	0.0	3.8	0.0	1.7	4.0	0.0	1.2	5.7	0.01	0.02	3.83
Range	9.6	9.5	5.7	9.7	7.8	5.5	9.7	8.3	3.0	9.65	9.53	5.72
DAISY												
Max	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3
3Q	7.3	7.3	7.3	7.3	7.25	7.3	7.3	7.3	7.3	7.3	7.3	7.3
2Q	7.1	7.1	7.1	7.1	7.1	7.1	7.1	7.1	7.1	7.1	7.1	7.1
1Q	6.9	6.9	6.9	6.9	6.9	6.9	6.9	6.8	6.4	6.9	6.9	6.9
Min	4.6	4.6	5.1	4.6	4.6	5.0	1.3	1.3	1.3	1.3	1.3	1.3
Range	3.7	3.7	3.2	3.7	3.7	3.3	7.0	7.0	7.0	7	7	7

Finally for all models, as expected, the median values of the distributions of simulated yields and simulated total growing season evapotranspiration of the Jülicher Börde region are higher than the corresponding values of the Sauerland region and only small interactions between model and region are simulated, i.e. simulated differences between regions are smaller for SIMPLACE<LINTUL-SLIM> than for the other models.

4.4 Discussion

4.4.1 Spatial aggregation of soil data

The low impact of soil data resolution on the distributions of simulated yields for the period and regions selected for this study can be attributed to three main reasons.

First, to calculate the distributions of simulated yields in each region we considered all years in the period from 1995 to 2008. Although no aggregation of results (averaging) was undertaken, the consideration of all years to calculate the probability distributions of simulated yields in a region might neglect the effect of the inter-annual variability of precipitation on the water balance. The heterogeneity of water retention properties of soils becomes an important yield-influencing factor when water supplied by precipitation is scarce (De Wit and van Keulen, 1987). Figure 4 shows exemplarily for the county of Aachen in the region Jülicher Börde a comparison between the distributions of yields simulated by all models for the driest (1996) and the wettest (2000) years of the selected period. The influence of the spatial resolution of soil input data on the density traces of simulated yields is not evident for the distributions of yields simulated for 2000. On the contrary, yield distributions simulated for 1996 remarkably differed among the tested spatial resolutions of soil data (Figure 21). In our study region for most of the years precipitation barely caused water stress in the model simulations. Hence, the distribution of soil properties for different spatial resolution has apparently not played a decisive role on average.

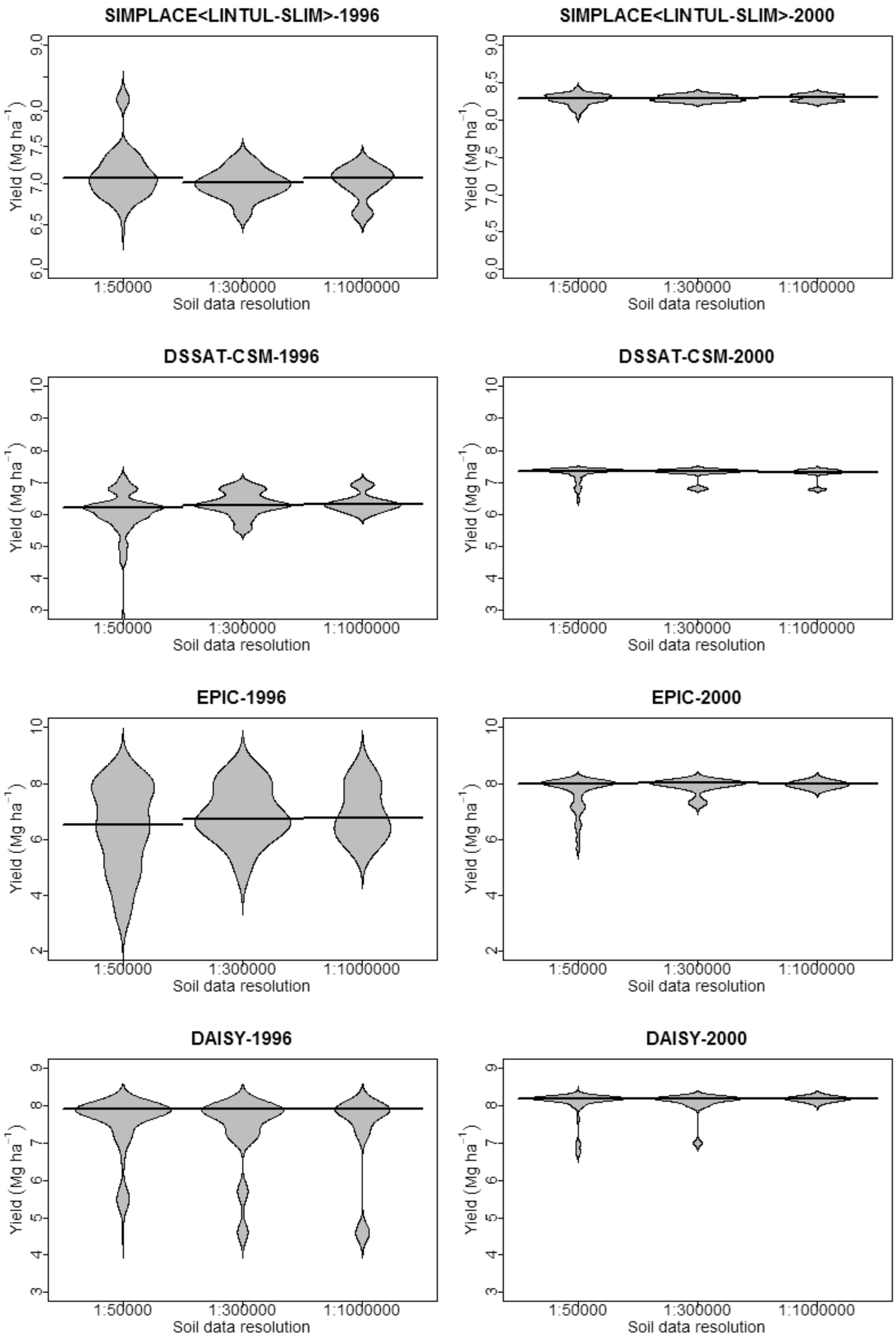

Figure 21. Comparison of frequency distributions of yields simulated by four crop growth models using three soil data resolutions for the county of Aachen for the driest (1996) and wettest (2000) years of the period from 1995 to 2008 (horizontal black lines in the bean plot represent the median value of the frequency distributions).

Second, we applied pedotransfer functions to estimate the water holding capacity of all used profiles and did not use any site-specific measured data on soil water holding capacity. This might have caused an artificial decrease in the variability of soil water holding capacity in our data set as compared to the actual variability of soil water holding capacity. In this respect, Lawless *et al.* (2008) undertook a numerical simulation analysis of the effect of the uncertainty emerging from using pedotransfer functions for estimating soil hydraulic properties on yield estimates in the UK. They concluded that the sole application of pedotransfer functions might be too coarse to estimate hydraulic soil properties used as input for mechanistic crop growth models. Consequently, they recommended combining pedotransfer functions with site specific soil water holding capacity measurements which might capture the spatial variability of hydraulic soil properties in a region. For large areas, however, such approach might not be feasible or hampered by the low density of individual measurements for large areas, which was the case in the present study.

Finally, the aggregation process to obtain the soil data for res2 and res3 can also be the cause of the very low impact of soil input data resolution on yield distributions. We selected the spatially dominant mapping units with their soil profile descriptions as representative soil units for the next coarser resolution (see 4.2.3). As a result, some profiles (i.e. the representative ones) and consequently the information in terms of soil water holding capacity values for the models were repeated in each simulation step for different soil input data resolutions. The method described above is the standard procedure used to generate soil maps with lower resolution, if high resolution maps are available, and is based largely on both formal knowledge and intuition (Heuvelink and Webster, 2001). Thus, the impact of soil data resolution on regional yield simulations might be influenced by the base resolution and the criteria for selecting representative soil units utilized to create the generalized regional soil maps which are used as input data for crop modelling applications. To get more insight into this matter, we undertook an additional test with one of the models. We analysed the distributions of yields simulated by SIMPLACE<LINTUL-SLIM> using the coarsest resolution (1:1 000 000) considering three criteria for selection of representative profiles: a) the best yielding profiles, b) the most surface dominant profiles as described in section 4.2.3, and c) the worst yielding profiles (Figure 22). The number of profiles selected in each county for the three mentioned criteria was given by the number of most representative profiles according to the description in section 4.2.3, for example for the county Aachen we built the bean plots of the yields of a) the 6 best yielding profiles, b) the yields of the 6 most representative profiles in terms of area and c) the 6 worst yielding profiles. Since the distributions of yields in all counties in the same region were very similar, we decided to present the yield distributions for the two regions and not for each county individually. As expected, the shapes and extends of the simulated yield distributions varied according to the criteria to choose the representative profiles (Figure 22). From these results it can be inferred that when the process of spatial aggregation of soil data uses the same criteria to select representative soil units for all aggregation steps, spatial resolution of soil input data apparently plays a negligible role as a source of uncertainty in regional crop growth simulation results. It may well be that when maps of different resolutions, produced with different aggregation procedures, are used, the soil input data may cause distinct frequency distributions depending on the resolution. It would be recommendable to undertake a study to prove this.

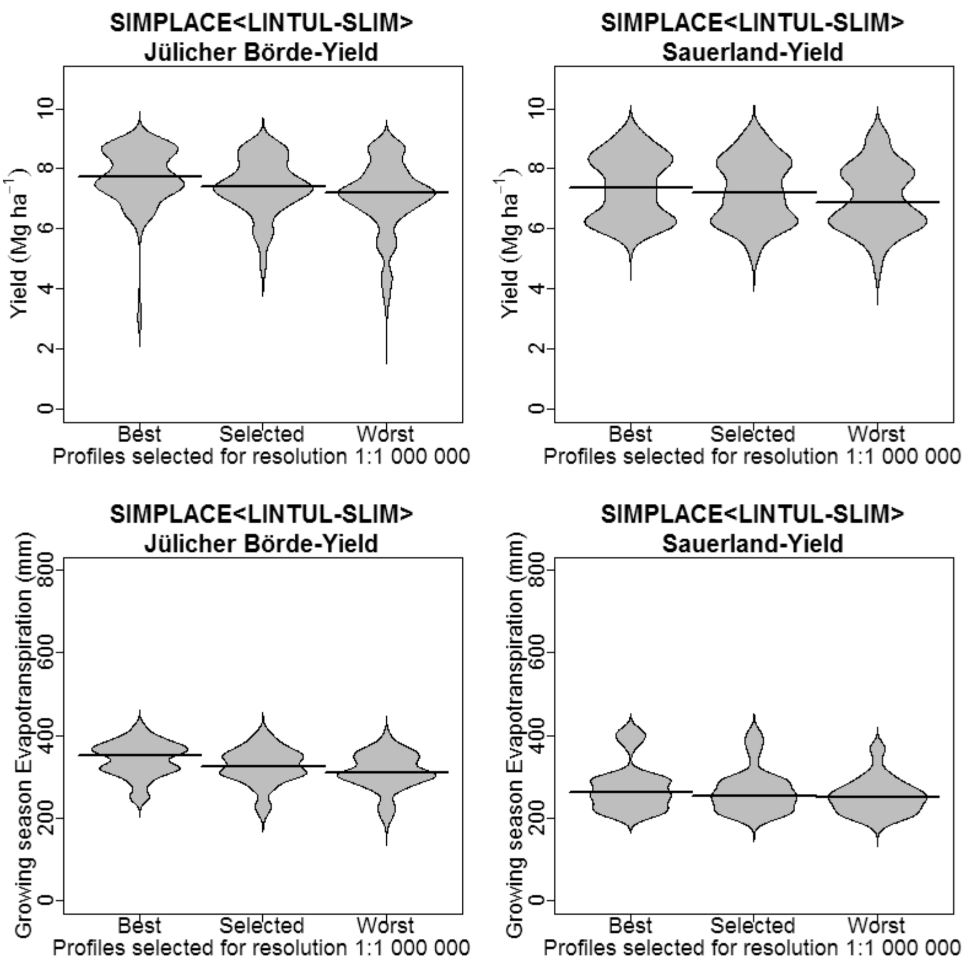

Figure 22. Comparison of frequency distributions of yield and total growing season evapotranspiration simulated by SIMPLACE<LINTUL-SLIM> considering the best soil profiles (left), the representative soil profiles as described in section 2.3 (middle) and the worst soil profiles (right).

4.4.2 Differences between models

The uncertainty in regional yield simulations caused by the model choice appeared to be larger than the uncertainty introduced by the resolution of soil input data for each model (Figure 20). We attempted a quantitative assessment in order to clarify the relative importance of both uncertainty sources. On the one hand, the coefficient of variation (CV) of the simulated results for each model and resolution was calculated. The CV offers a summary description of the variability of simulated yields and was used as a measurement of the uncertainty introduced by the model choice. On the other hand, for assessing the uncertainty (error) introduced by the use of coarser resolutions of soil input data, we calculated the coefficient of variation of the root mean square error (CV(RMSE)) of the yields simulated with res2 (1 : 300 000) and res3 (1 : 1 000 000) compared to res1 (1 : 50 000 highest resolution). Figure 23 offers an overview of CV and CV(RMSE) values for all models and resolutions.

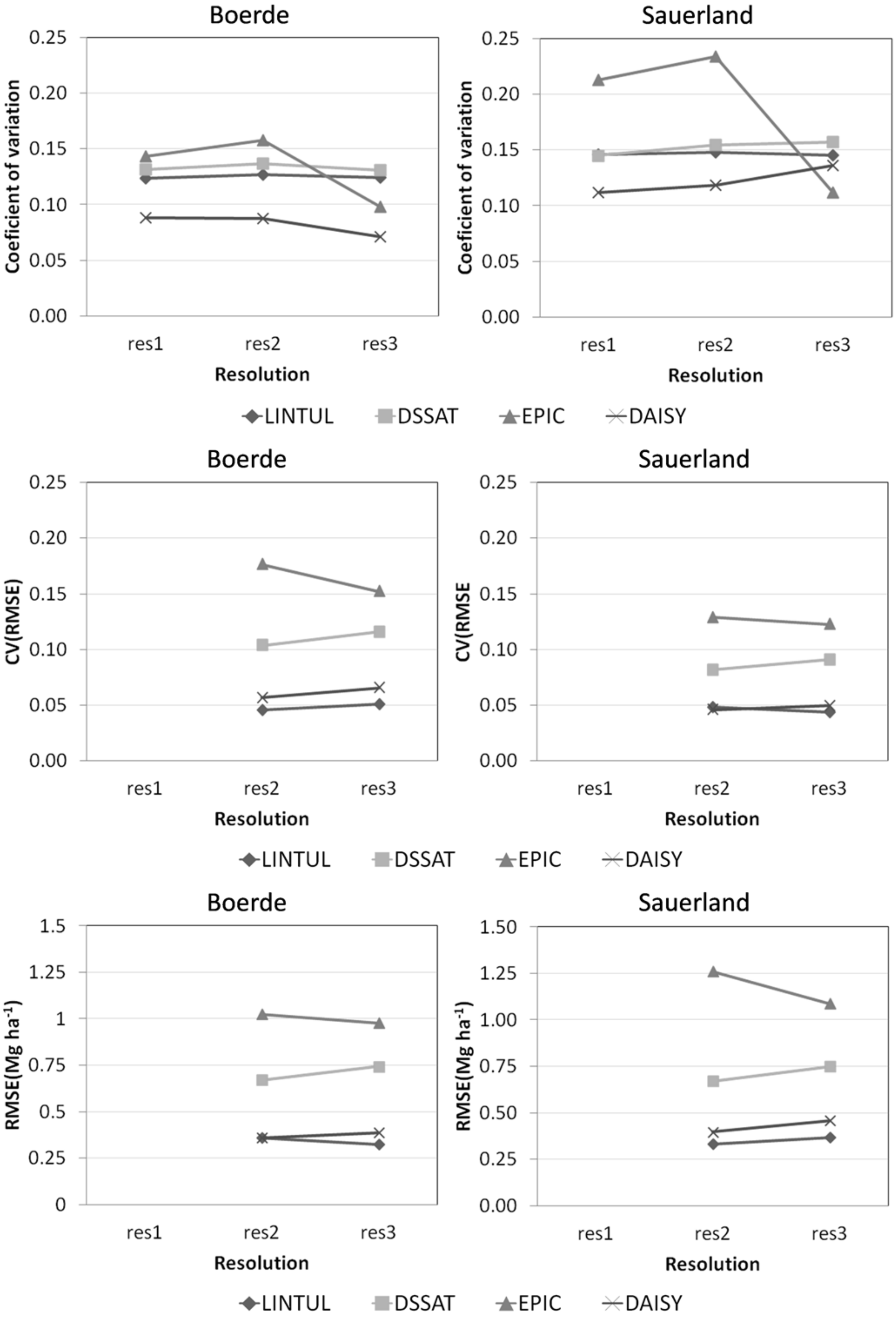

Figure 23. Comparison of coefficient of variation (CV), coefficient of variation of root mean square error (CV(RMSE)) and absolute root mean square error (RMSE) of the simulated crop yields of four crop growth models using three soil data resolutions. (For RMSE calculations simulated crop yields in res1 are considered to be the best approximation to the observed yields)

One would expect that using a lower i.e. coarser resolution of soil input data might cause a decrease in the variability of simulated yields. In both considered regions this was only partially the case for the model EPIC. Its CV values decreased from 0.15 (res2) to 0.09 (res3) in Jülicher Boerde and from 0.23 (res2) to 0.11 (res3) in Sauerland. However, for all other models the CV values either increased or decreased only very slightly (first panel in Figure 23). Concerning the error introduced by a coarser resolution, it could be assumed that the values of CV(RMSE) for res3 (1 : 1 000 000) would be higher than the ones for res2 (1 : 300 000), since in general, lower resolution maps contain less detailed information. However, similarly to CV, the values of CV(RMSE) did not show any ascending trend when lower resolution soil input data were utilized. The least difference between CV(RMSE) of res2 and CV(RMSE) of res3 was 0%, thus, no difference for SIMPLACE<LINTUL-SLIM> in the region Jülicher Boerde. The highest difference of 3% was found for the Model EPIC in the region Jülicher Boerde.

In contrast, the range of differences between models when comparing the values of CV of yields simulated at one resolution varied between 1 and 11%. Likewise, the differences of values of CV(RMSE) between models at one resolution ranged between 1 to 13% (second panel in Figure 23).

When considering the absolute values of RMSE (third panel in Figure 23), it becomes even more evident that the error introduced by the choice of the model, given by the ranges of RMSE between 0.33 to 1.26 Mg ha^{-1} is greater than the uncertainty caused by the resolution of input data in each model which ranged between 0.07 Mg ha^{-1} for DSSAT-CSM in the Jülicher Boerde region to 0.17 Mg ha^{-1} for EPIC in the Sauerland region.

4.4.3 Possible causes of model differences

The differences in terms of shape and extent of simulated yield distributions between the four models (Figure 19 and Figure 20) can be basically attributed to specific structure and implementation of each model (Adam et al., 2012; Angulo et al., 2013b; Mearns et al., 1999). A detailed causal analysis of these differences is beyond the scope of the present study; however, based on our results, we provide some points for discussion which may stimulate more advanced analysis and research in the future.

Crop yield is the result of linear and non-linear interactions between environment, genotype and management. Therefore, it was not expected to find a clear linear relationship between simulated yields and simulated total growing season evapotranspiration. However, as we avoided the influence of management practices on yield simulations by using the best management practices recommended for our study region, we expected to find some correspondence between simulated yield variability and total growing season evapotranspiration, which was not the case. This lack of correspondence could be attributed to the differing degree of importance by which every model considers soil water dynamics calculations as determinant of simulated yields. For instance, in the case of DAISY the relative differences (1) of the distribution median values between Jülicher Börde and Sauerland in each resolution are almost equal for simulated yield (res1: 0.1, res2: 0.09 and res3: 0.10) and simulated total growing season evapotranspiration (0.09 for all resolutions). Contrarily, for SIMPLACE<LINTUL-SLIM>, the relation of the relative differences of the median distribution values between Jülicher Börde and Sauerland is approximately 1 to 5 for simulated yield: res1: 0.04, res2: 0.04, res3: 0.03 and for simulated total growing season evapotranspiration res1: 0.21, res2: 0.22, res3: 0.22.

$$\text{relative difference} = \text{med Jülicher Börde}_n - \text{med Sauerland}_n / \text{med Jülicher Börde}_n \qquad (1)$$

where : n=resolution; med=median value of distribution

Despite our assumption to exclude management effects from the analysis, we could not exclude the influence of model yield reducing factors which could also have had an influence on simulated yields. A clear example of the described phenomena is DSSAT-CSM. When analysing the probability distributions of simulated temperature stress of DSSAT-CSM for the two considered BKRs (not shown) more pronounced temperature stress levels were found for Sauerland. This can be explained by the lower winter temperatures in the Sauerland region in comparison to the Jülicher Börde region (Table 11). Thus, for DSSAT-CSM the differences in simulated yield levels between Jülicher Börde (BKR141) and Sauerland (BKR134) appear to be caused mainly by differences in temperature and not by soil water dynamics i.e. simulated total growing season evapotranspiration.

For the selected regions and period it was not possible to find single ‘fingerprints’ with which a model might be identified. The shapes of the simulated yield and total growing season evapotranspiration by one model were similar for the same region but differed between regions. However, in agreement with the conclusions of earlier studies (Angulo et al., 2013b; Willmott, 1981; Willmott et al., 1985), it appears to be recommendable to evaluate the simulation results of crop models regarding whole distributions rather than focusing only on summary statistics when one is interested on assessing the spatial and temporal (year to year) variability of regional yield simulations. For example, the median values of the distributions of yields simulated by EPIC and DSSAT-CSM for the Jülicher Börde region are very similar. Nevertheless, by assessing visually whole distributions depicted in the form of bean plots (fingerprints), it is possible to distinguish that for the selected period the highest density of simulated yields is between 8 and 9 Mg ha^{-1} for EPIC while for DSSAT-CSM the simulated yields spread relatively uniformly between 7 and 10 Mg ha^{-1}.

4.5 Conclusion

In this study we show that for the selected regions, period and models the choice of soil input data resolution has little influence on the shape and extent of the distribution of simulated yields and total growing season evapotranspiration. We identify three reasons for this response: a) the high precipitation amounts in the region which diminish the importance of soil-dependent water supply; b) the loss of variability of the hydraulic soil properties related to the methods applied to calculate water retention properties of the used soil profiles; and c) the method of aggregation of soil data used in our study, which considered the same representativeness criteria for the three resolution levels based on the same soil profile database for all aggregation steps. Further research on evaluating the distributions of crop model regional yield simulations for different soil data spatial resolutions might explicitly consider different aggregation methods. For assessing the behaviour of various crop models, when different soil input data resolutions are used for simulating regional yields, it is recommendable to evaluate the model results as whole distributions, if one is interested in a fast and clear assessment of the year to year variability of regional yield distributions. Since in our study the form and partially extent of each individual model fingerprint depends not only on the model but also on the interaction between inter-annual weather variability and soil properties, it might be advisable to undertake further evaluation considering the interactions between soil and weather input data resolution in order to clarify the applicability of the 'fingerprints' as model typifying tool. Thus, the use of fingerprints at the moment is limited to offering a qualitative estimate into the temporal and spatial variability of simulated regional yields. According to the results of the present study, the uncertainty introduced by the model choice seems to be more important than the uncertainties caused by the soil input data resolution. Therefore, we suggest applying a multi-model ensemble approach to regional studies including the assessment of the effect of different scaling methods.

Chapter 5

General Discussion

5. General Discussion and Conclusions

This PhD thesis was carried out in response to the urgent need of developing approaches for identifying, quantifying, reporting and (ideally) reducing the uncertainty emerging from the regional application of field scale crop models, particularly in the context of climate change impact studies (Asseng et al., 2013; Rosenzweig et al., 2013; Rötter et al., 2011a). This chapter discusses the main outcomes of the research presented in chapters **2** to **4**. The first section of this chapter (section 5.1) addresses methodological issues of the presented analyses with particular emphasis on the assessment of uncertainty and the scaling up of crop models. Furthermore, the innovative solutions and shortcomings of the studies as well as new research questions emerging while answering the original research questions of the thesis introduced in chapter **1** (Q1-Q3) are presented in the four following sections (5.2 to 5.5). Finally, section 5.6 presents the general conclusions on how to manage data uncertainty in regional crop model applications.

5.1 Methodological issues to characterize uncertainty in regional crop model applications

5.1.1 Simulation experiments

This thesis paid special attention to **systematically** analyse the effects of the spatial resolution of weather and soil input data (Q2, Q3) and different calibration strategies (Q1) on the uncertainty of regional crop model simulations. Accordingly, for studying Q2 and Q3, the spatial resolution of weather and soil input data was systematically reduced (from higher to lower resolution) within a range of 10 km x 10 km to 100 km x 100 km grids for weather data and maps of scale 1:50000 to 1:1000000 for soil data using the same data basis for all resolutions. The three calibration strategies studied in chapter 2 represent different degrees of complexity in calibrating crop models from relatively simple (using only phenology related parameters) to a rather elaborated strategy (using also crop growth related parameters) (Q1- section 2.2.5.2). However, the choice of study regions, crops and crop models used did not follow a strictly systematic approach, but was the result of data availability in the study region and the ability of modelling groups to engage in this study which is also explain further in the following sections.

Selected study regions and crops

Since the studies presented in chapter 2 to 4 were highly dependent on the quality and quantity of input and validation data, the decision of the regions and crops to study was mainly taken based on the availability of data (see 1.4). The reader might wonder why three different spatial extents were considered for the three proposed research questions: continental scale for Q1, basin scale for Q2 and sub-national for Q3. The main objective of Q1 was to test whether region-specific parameterisation as proposed by earlier studies in Europe (Reidsma et al., 2009a; Reidsma et al., 2009b; Van Der Velde et al., 2009) improved the simulation results. Here, the continental scale (Europe (EU25) appeared to be most adequate for this purpose as crop varieties differ across Europe and as good and accessible data were available. In Q2 and Q3 whole Europe was not taken into consideration for various reasons. First, the available resolution of weather and soil input data was too coarse to systematically assess the uncertainty introduced by different spatial resolutions. Moreover, most regional climate change impact projections on crop production are performed at (sub-)national scale, and finally, the time constraints for a systematic scaling study from plot to European level would have been beyond the scope of a single PhD study. In chapter 2 five major crops where considered to investigate Q1. Winter wheat, winter barley, potatoes, sugar beet and maize are very important crops in Europe. Although crops such as rape seed, spring wheat and silage maize are also important, cultivation and yield data for the selected five met best the criteria to be used in our study. Ideally, we should have considered the same cereal crop, for the studies in chapters 3 and 4. However, due to the differences between study regions, the most representative crops, spring barley and winter wheat were taken into consideration for the study cases in Finland (Yläneenjoki region) and Germany (North-Rhine Westphalia), respectively. Despite the obvious differences between both crops, such as the length of growing period and vernalisation requirements, results on the effects of the resolution of weather input data for both crops were similar (Figure 13 *vs.* Figure 16).

Selected models

The crop model LINTUL2 (Angulo et al., 2013a) was used in order to answer Q1. The choice was based on the relative simplicity of the model, and the working experience of our research group (INRES-Crop Science, University of Bonn) with it. In the case of Q2 and Q3, the most important criteria to choose the crop models were to have models of different complexity and structure and research groups competent in applying them. Final choices were strongly co-determined by close research contact with crop modelling research groups having ample working experience with the respective models and the capacities to carry out the required simulations. Crop models run by collaborating research groups were DSSAT-CSM (Jones et al., 2003a), DAISY(Gassman et al., 2004), and WOFOST 7.1 (Boogaard et al., 1998; Van Diepen et al., 1989). These differ structurally distinctly from the two other models LINTUL2 and EPIC run at INRES-Crop Science, University of Bonn. Special attention was paid to the model differences regarding detail of light interception and light utilization for biomass assimilation processes (see Adam et al., 2011).

Furthermore, it might have been interesting for the study undertaken in Chapter 3 (Q2) to consider a more detailed model of soil water dynamics such as the Richards approach in comparison to models using the conventional tipping bucket approach. This was not possible for Q2 due to time constraints. However, in the study presented in Chapter 4 (Q3), a model using the Richards approach, DAISY (Hansen et al., 2012), was included in the model comparison exercise.

Why focus on water-limited yields?

In all the studies of this thesis simulations refer to water-limited yields. In order to answer Q1 it was assumed that temperature and water supply might be the most important factors explaining yield variability in Europe. Thus, the parameters selected to test calibration strategy 3 (Q1-section 2.2.5.2) were assumed to represent the differences in the varieties grown between sub-regions to cope with water deficiency. Other yield-limiting factors (nutrient deficiencies, pest and diseases) were not explicitly modelled but were likely inherently captured when calculating the growth parameters.

For the studies in chapters 3 and 4 (Q2, Q3) simulations for rain-fed conditions appear justified since the selected study regions, West Germany and South-West Finland, are production zones where farmers generally apply ample fertilizer and crop protection to achieve actual yields under the given rainfall regimes that are close to yield potential. Other factors influencing productivity indirectly through crop management such as agricultural and environmental policies and market regulations are not considered in crop growth models but in agro-economic models and were therefore not considered in our studies. As crop management in the regions investigated in chapters 3 and 4 is typically close to optimal simulation of water-limited yields seemed sufficient to gain first insights into effects of scaling on regional yields.

5.1.2 Considered scaling method

The focus of the research efforts undertaken in chapters 3 and 4 (Q2, Q3) was to explore the uncertainty introduced into regional crop model simulation results by **spatial scaling of input data**. The specific scaling method investigated was aggregation (for a definition see section 1.1.1).

Weather data

Aggregation was selected since it is commonly applied as a strategy to spatially scale weather and soil input data for regional crop modelling (e.g. De Wit et al., 2010; De Wit et al., 2005; Easterling et al., 1998; Ewert et al., 2011b; Folberth et al., 2012; Mearns et al., 2001; Mearns et al., 2003; Mearns et al., 1999; Nendel et al., 2013; Niu et al., 2009; Reidsma et al., 2009a; Rötter et al., 2013b; Rötter et al., 2011b; Van Bussel et al., 2011a). Weather input data are obtainable in form of grid cells for a number of regions and the availability of gridded weather data for regional analysis is increasing (Folberth et al., 2012). Often the results of weather generators from downscaling the results of GCM outputs are used for both baseline and future weather conditions as inputs for regional climate change impact assessments (Semenov and Pilkington-Bennett, 2012; Semenov et al., 2013). Such data are also available and used as gridded data. Thus, depending on the scale considered in a regional yield assessment study, a weather grid cells size, i.e. a specific aggregation level of gridded weather data is often used. In this respect, systematic studies of the resolution/grid cell size requirements for assessing climate change impacts on crop yield are missing. Accordingly, three aggregation levels (4 spatial resolutions) were tested: starting from a base weather data grid of 10 km x 10 km, which was step-wise aggregated to 20 km x 20 km, 50 km x 50 km and 100 km x 100 km grids. The additional consideration of data from a near weather station in chapter 3 served to illustrate the differences between point and gridded weather input data.

Soil data

In general, soil profiles or, more frequently, soil maps based on soil surveys are the primary source of generating soil input data for regional crop model applications (Bechini et al., 2003). In many countries soil information at high resolution (mapping scale higher than 1:200000) is scarce due to the extremely high cost of soil mapping. Therefore, it is important to quantify the minimum resolution which is necessary to adequately reproduce crop yields in the context of regional crop modelling applications. The commonly used technique to generate soil maps (i.e. generalization in cartographic terms) from higher to lower resolution, is to unite (larger) areas that share similar physical/geological characteristics (soil sub-units) into a generalized class characterized by a typical soil profile (Leenhardt et al., 1994). Thus, choosing the scale of a soil map, to be used as source of soil input data means in practical terms to choose a specific spatial resolution which may or may not be supported by the underlying point measurement or support data (such as geo-morphological boundaries derived from diverse sources). Accordingly, the study in chapter 4 (Q3) tested the effect of two aggregation levels (three soil map resolutions) on regional yield simulations.

Since the effect of aggregation depends on the properties of the utilised data (Van Bussel et al., 2011a), the obtained results are in the first place only applicable for regions were the data situation and the weather and soil conditions and their heterogeneity/homogeneity reflected in the data are similar to our studies (see sections 5.2, 5.3).

Data quality

All data from model simulations and observations are prone to error also as a result of data manipulation through scaling such as aggregation. According to the results presented in chapter 3, aggregating observed yields has a distorting effect on the shape and range of the probability distribution of the data (Q2-Figure 15). This has repercussions on the interpretation of observed data used to calibrate crop models. A first attempt to visualize the mentioned effect of aggregation, for example simple averaging is presented in Figure 24. On the one hand, for winter wheat, a crop for which many observed yield data were available, the aggregation of observed and simulated yields has practically no effect on the shape but it shortens the range of the frequency distributions. On the other hand, for sugar beet, a crop for which data on observed yields were relatively scarce (30% less data in comparison to winter wheat), aggregation led to a levelling of the shape and reduced the range of the frequency distributions of observed and simulated yields. This shows that the impact of aggregation as well as the success of a calibration strategy is also closely related to the quality and /or quantity of base data (Hansen and Jones, 2000).

The quality of model input data is essential prerequisite for good model performance. Available weather data have been provided by the DWD for the study in chapter 4 after initial quality check. However, some data were not sufficiently available such as wind speed and assumptions had to be made which may have had some implications on the obtained results (see section 5.2).

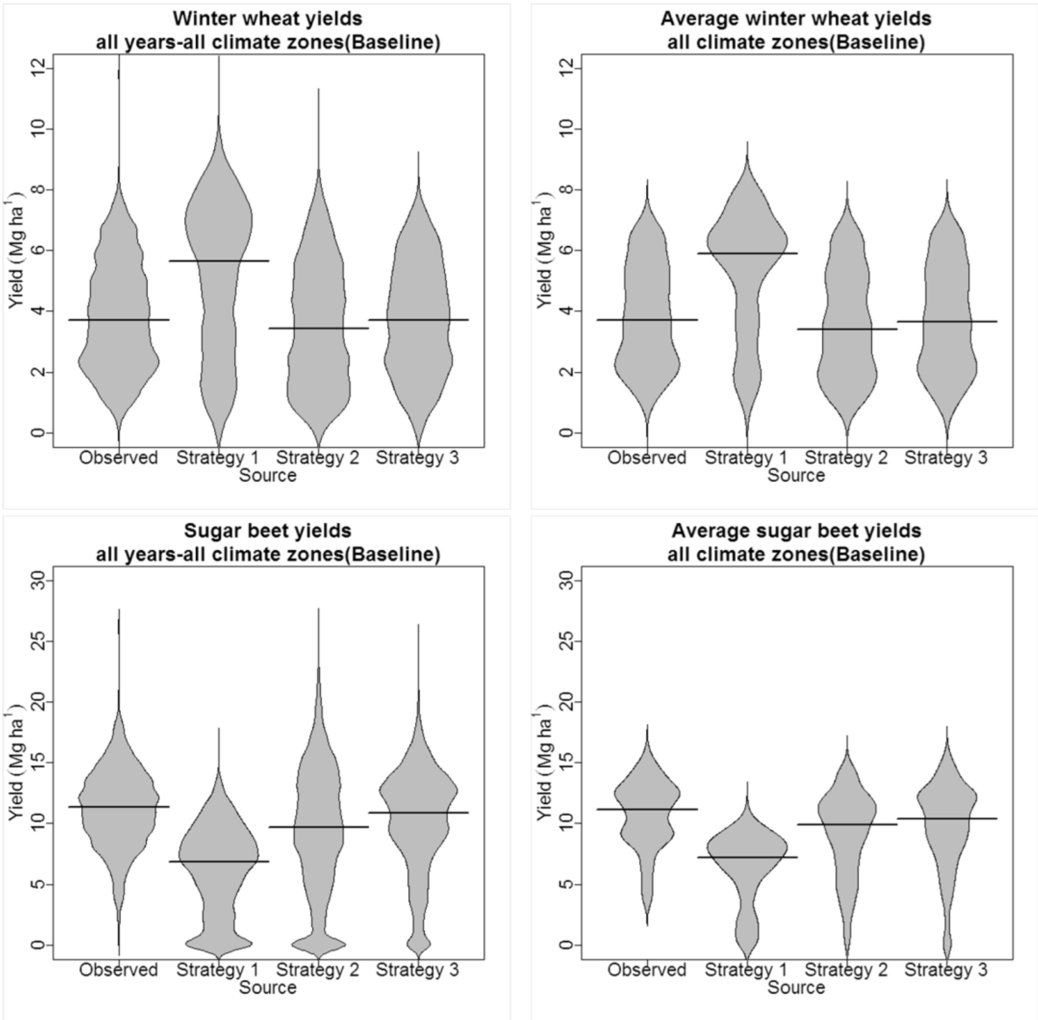

Figure 24. Interacting effect of aggregation of simulated yields of winter wheat and sugar beet simulated for 533 climate zones in the period of 1983 to 2006 in relation to three model calibration strategies. Strategy 1: phenology only, Strategy 2: using a yield correction factor, and Strategy 3: extended calibration of selected growth parameters of winter wheat for 533 climate zones in Europe in the period from 1983 to 2006. See text for explanation of calibration strategies. Left panels show probability distributions built up considering all years and all climate zones for winter wheat (above) and sugar beet (below). Right panels show probability distributions built up with the average values over years of each considered climate zone.

5.1.3 Evaluating uncertainty

Model evaluation

In the present study, uncertainty in regional yield simulations was assessed from two different viewpoints. On the one hand, chapters 3 and 4 (Q2, Q3) addressed the effect of the spatial resolution of input data on model simulations. The yield predictive power of the models was assumed to be plausible and was not further evaluated and compared with observed data. On the other hand, chapter 2 (Q1) evaluated the discrepancies between observed and simulated spatial yield variability as measurement of the uncertainty introduced by different parameter estimation strategies (section 5.4 will discuss variability depiction as indicator of plausible model results).

Visualisation of results

Although the effects of soil and/or weather input data resolution have been partially investigated in several studies (e.g. Folberth et al., 2012; Mearns et al., 2001; Moen et al., 1994; Nendel et al., 2013; Niu et al., 2009; Wassenaar et al., 1999), none of these studies applied a systematic approach to explicitly evaluate the uncertainty in crop model results caused by: i) input data resolution, ii) model structure and iii) the interaction of both. In support of (gradually) closing this knowledge gap, the use of bean plots was chosen as a major tool to illustrate the variability addressed in Q2 and Q3.

Similarly to box and whisker plots, bean plots depict the degree of dispersion and skewness of a data set and do not make any assumption about the statistical distribution to which the data might correspond. An additional and very advantageous feature of the bean plots for our work was their ability to estimate the frequency distribution of the data sets (Sheather, 2004; Sheather and Jones, 1991; Silverman, 1986), represented by the outer form of the beans (e.g Q2-Figure 13).

For the evaluation of simulation models it has been suggested not to rely only on the correlation coefficient and its square, i.e. the coefficient of determination, but to use various other statistical measures for characterizing model performance (Willmott, 1981; Willmott et al., 1985). Increasingly, crop modelling studies have been following these recommendations and made use of statistics such as root mean square error (RMSE) and model efficiency (ME) (e.g. Nendel et al., 2013; Palosuo et al., 2011; Rötter et al., 2012b). In this regard, the present PhD thesis is innovative for being the first work in the area of crop modelling using whole distributions in the form of bean plots to evaluate model outcomes.

Due to the high amount of graphical information offered by the beans, a rapid visual assessment is possible to gain qualitative insight into the variability in the data analysed. For instance, the similarity in shape and extension between all bean plots of yields simulated by the same model but using different weather data resolutions (Q2-Figure 13) underlines the low effect of weather data resolution on uncertainty of model results at the selected sites. An additional advantage of using bean plots is the low cost of implementation. Bean plots are developed in the free software R (*http://CRAN.R-project.org/package=beanplot*).

However, bean plots also have constraints. Although they allow visual assessment of distributions, no quantitative information is provided. It is therefore necessary to extend the analysis into more formal statistics. Bean plots also do not provide explanation of different distributions. Thus, further analysis of causal relationships is required to explain the remarkable differences between the frequency distributions of the different crop models as found in this thesis. Bean plots are also not very useful when the predictive power of a single model is to be validated. Nevertheless, depicting observed yields for a specific region and time period in form of frequency distributions offers a rough estimate of the observed temporal and spatial regional yield variability. In theory, the bean plots of yields simulated by a well calibrated crop model should be similar to the bean plots of observed yields (see section 5.1.2).

5.2 Influence of weather input data resolution on simulated yields

According to the results of chapter 3, changing the resolution of weather data does not markedly increase the uncertainty of crop model simulations (Q2-Figure 13). However, this statement only applies to regions where the properties of weather are similar to the study region under consideration in chapter 3. The Yläneenjoki region in Southwest Finland is characterized by a very homogeneous topography, which in general terms implies very small differences in temperature and precipitation between sub-regions (Johansson and Chen, 2003). Thus, although temperature differences between grid cells do exist, they did not appear to be large enough to make aggregation (scale change) significantly influencing the yield simulations. In concordance with these results, it has been shown that in regions where temperature values are similar between sub-regions, a finer resolution does not improve the simulation results of crop phenology (Van Bussel et al., 2011a). Moreover, aggregation of precipitation data did not have any remarkable influence on the variability of simulated yields or total evapotranspiration during the growing season. It has been suggested that the aggregation of precipitation data in grid cells might cause an artificially homogeneous daily distribution of the daily amount of water supplied to the plant (Hansen and Jones, 2000). From our results it can be inferred that the variability lost caused by aggregating precipitation data should be considered only when the differences in precipitation between sub-regions surpass a certain threshold. The determination of this precipitation variability threshold is still to be investigated and might be region specific. Eventually, it will be necessary to undertake a similar systematic study in a study region characterized by less homogeneous topography.

One limitation to test the effect of weather data aggregation on the results of regional crop model applications was the already mentioned lack of gridded wind speed data. In our case study one data set for wind speed was used for all resolutions. Three of the four models (Q2-Table 3) calculated the needed values of potential evapotranspiration using equations which require wind speed as input parameter (Allen et al., 1998a; Monteith and Greenwood, 1986; Penman, 1956). Since the values of evapotranspiration are a crucial internal variable to calculate water dynamics in all the models, it could be argued that the very low impact of weather data aggregation on model simulations might also be related to the fact that similar potential evapotranspiration values served as a base for water stress calculations in all aggregation levels for each model. However, when comparing the results of DSSAT-CERES, which uses an equation not requiring wind speed (Priestley and Taylor, 1972) with the remaining three models, no apparent difference in terms of variability introduction can be noticed (Q2-Figure 13). Yet, this might also be a region-specific phenomenon, as the aerodynamic term in the evapotranspiration estimation usually plays a minor role in humid, temperate or boreal climates as in Finland. The results point out the need for further research to quantify the influence of the methods to calculate evapotranspiration on regional yield simulations. If methods not considering wind speed cause a similar uncertainty as methods using it, crop models may use routines that do not require wind speed (at least in certain regions) and hence might overcome the problem of limited availability of wind speed data.

Although previous research has already considered the influence of different weather data resolutions on regional yield simulations (Folberth et al., 2012; Mearns et al., 2001; Nendel et al., 2013; Olesen et al., 2000) the work described in chapter 3 is the first systematic approach using the same data basis for all resolutions (aggregation steps). The small influence of weather data aggregation found in our study might partly be due to the small error introduced when using the same data basis for all resolution steps as compared to studies with diverse databases.

Quantification and reporting of uncertainty in regional crop model applications require transparency in the processes of obtaining and processing input data. Therefore, there is an urgent need for the crop modelling community to search for cooperation in order to collect weather data sets as needed for crop model applications. It is recommended to involve experts in meteorology/climatology when processing, interpreting, and scaling weather data.

Gaining a better understanding of the influence of scaling weather data for regional crop model applications might facilitate the choice of the most appropriate resolution needed when utilizing (regional or global) climate models for climate impact assessments (Mearns et al., 2001; Semenov and Pilkington-Bennett, 2012; Semenov et al., 2013; Semenov and Shewry, 2011). For instance, in regions like the Yläneejoki region, where topography is fairly homogeneous, the effort of downscaling weather data to higher resolutions might not represent a gain in the quality of yield simulation results. From the results of our study it can also be recommended, that for regions where weather variables such as temperature and precipitation are evenly distributed, methods such as sampling of a representative grid cell or weather station might be enough to represent the weather variability of the region.

The influence of rainfall on crop growth is closely related to the water retention characteristics of soil. Since model simulation results are usually influenced by the interaction between precipitation and soil characteristics affecting soil water availability (e.g. Folberth et al., 2012; Nendel et al., 2013; Niu et al., 2009), the results discussed in chapter 3 depict only partially the effect of aggregation of weather data on regional yield simulations. In the next paragraph results on the effect of aggregating soil characteristics are discussed.

5.3 Effects of soil data aggregation

Similarly to the results of chapter 3, the results of chapter 4 also reveal that the spatial aggregation of soil input data did not have a considerable effect on the variability of crop model yield simulations (Q3-Figure 19). Motivated by the experience gathered in chapter 3, a topographically more heterogeneous region was chosen to answer Q3. Although noticeable differences in terms of simulated yield and simulated total growing season evapotranspiration distributions between the plains and the mountainous sub-regions were apparent (Q3-Figure 19), no considerable impact of soil data resolutions within a sub-region was found.

Most of the studies which have investigated the impact of soil data resolution on regional yield simulations have not focused explicitly on the uncertainty introduced by a specific form of scale change (Easterling et al., 1998; Folberth et al., 2012; Nendel et al., 2013; Niu et al., 2009; Olesen et al., 2000; Wassenaar et al., 1999). An approach considering scaling systematically has been hampered by the lack of extensive and reliable soil data sets. Although the data basis in the region investigated in chapter 4 is exceptionally good, there is still a remarkable room for data quality improvement. For instance, the soil data used for the study were not collected by an agricultural service but by a geological service. Very important soil information such as soil texture, gravel content and soil water table was available in a well-documented data base. Nevertheless, the water contents at field capacity and wilting point were not provided per horizon but per profile. This water content at field capacity was estimated at a water tension of 0.06 MPa instead of 0.33MPa which is required by the crop models used. The minimum data requirements of all crop models to simulate soil water content are wilting point, field capacity and saturation in each soil layer which were estimated based on pedotransfer functions developed by the German soils (AG-Boden, 2005). Even if quality control based on expert knowledge was undertaken, the utilization of pedotransfer rules or functions introduced additional uncertainty. However, since the study in chapter 4 focused on the mere influence of scale change, the uncertainty introduced by the utilization of pedotransfer functions was assumed to be equal for all resolution levels.

It has also been shown that the utilization of site specific measurements of soil hydraulic properties to validate pedotransfer functions yields plausible values to be used as model input (Lawless et al., 2008). This underlines the urgent need for cooperation between soil scientists, geologists and agronomists to assure effective soil data collection with available resources.

An important source of uncertainty, which is usually neglected, is the criteria taken into consideration to select representative soil units. A clear example is given by the results presented in chapter 4 (Q3-Figure 22). The three proposed criteria for selecting representative soil units (1. high yielding profiles, 2. the most representative profiles in terms of area and 3. the worst yielding profiles) have a remarkable influence on the shape and range of the distributions of simulated yields and simulated total growing season evapotranspiration. Therefore, it would be interesting to compare soil maps of a specific region depicting the same or similar resolution but produced by different institutions/research groups applying different criteria to select representative soil units. Such study might offer a powerful insight into the so called "human error" caused by different representativeness-assumptions made in the process of soil data aggregation.

In chapters 3 and 4 it was intended to assess individually the effects of spatial aggregation of weather input data and the effects of spatial aggregation of soil input data, respectively. In both cases, no uncertainty introduction in simulation results was apparent. However, when considering the interaction between the year to year variability of precipitation and the soil characteristics, clear differences between soil resolutions were found (Q3-Figure 21). Such results recommend carrying out studies using a factorial simultaneous analysis of the influence of both weather and soil input data resolution on regional crop simulations.

5.4 Is it necessary to consider spatial heterogeneity in the model calibration process?

The parameters of plot/field crop models typically refer to crop growth and development processes and therefore are valid only for the scale at which they were developed (Challinor et al., 2009a). Due to the sub-regional differences of the factors affecting crop yields such as weather and crop management determined by farm characteristics, technology development and socio-economic conditions (Reidsma et al., 2009b), it is crucial for the calibration of models used in regional applications that model parameters reflect the spatial variability of such yield influencing factors (Hansen and Jones, 2000; Jagtap and Jones, 2002; Therond et al., 2011; Xiong et al., 2008). However, the majority of studies undertaken in the context of regional crop model application have not tested calibration strategies to solve this important issue (e.g. De Wit et al., 2010; Harrison and Butterfield, 1996; Van Der Velde et al., 2009). Based on the recommendations of an integrative European crop modelling study (Therond et al., 2011), chapter 2 (Q1) investigated the importance of region specific parameters to reproduce spatial heterogeneity in crop yields. The study did not pursue to develop a standard crop model calibration methodology but to compare calibration strategies about which and how parameters should be estimated (see 2.2.5.2). The third calibration strategy tested in this study, i.e. taking into consideration sub-regional differences of model parameters related to crop growth in addition to crop phenology resulted in the best agreement between simulated and observed yield at the European scale (EU25) (Q1-Figure 1). Nevertheless, since accurate calibration of crop growth and development parameters requires data which are presently scarce in the required quality and resolution for entire Europe, the use of a yield correction factor after phenology calibration (strategy 2) might be still meaningful and is advised as the preferred strategy.

Clearly, our results stress the need to consider uncertainty due to calibration as integrated part of the general reporting of uncertainty by the crop modelling community (Asseng et al., 2013; Rötter et al., 2013b). It should thus also be part of a common protocol to assess uncertainty in regional crop modelling applications as proposed for AgMIP (Rosenzweig et al., 2013) and MACSUR (Rötter et al., 2013a).

The results in chapter 3 showing an effect of aggregation on the shape of the probability distribution of the observed yield (Q2-Figure 15) have implications for the way of interpreting the results of chapter 2. If the model calibration process limits to "fit" the average results of our crop models to the average of observations (Challinor et al., 2009a), simulation results might be misleading and not usable in the broader context of decision making. This leads back to the issue of data quality which has been already mentioned in sections 5.2 and 5.3 and needs to be tackled immediately by the global crop modelling community. An effective way of solving this issue is the establishment of data transfer protocols specifically designed for crop model calibration in a multilateral frame such as in the European FACCE JPI project MACSUR (www.macsur.eu, Rötter et al., 2013a).

Finally, the utilization of easy to understand means such as graphics to present and discuss the results in chapter 2 considerably facilitated the collaborative work with other scientists such as economists involved in the project AgriAdapt (Ewert et al., 2011a) under which the study was performed. Clearly, impact assessment work is only possible when the product of the research by the crop modelling community is also understandable and usable for other scientists, decision makers and stakeholders which may need more attention in the future.

5.5 Next research steps on scaling methods

As a result of the work on scaling issues related to input data and model calibration a number of new research questions have emerged.

Although at first sight the spatial resolution of weather and soil input data alone does not seem to have an important impact on the frequency distributions of simulation results of regional crop models, it is recommended to undertake similar studies in regions where weather and soil data are more heterogeneous and the occurrence of drought stress is more frequent. It would be interesting to search for the feasibility of analysing the effect of soil data resolution in the Jokioinen river basin as well as investigating the effect of weather data resolution in North-Rhine Westphalia. In this regard, a systematic factorial analysis of the combined effect of soil and weather data resolution on simulated yield variability might bring more insight into the effect and related uncertainty of input data resolution for regional crop modelling. Additionally, it might be interesting to systematically analyse the interaction between the year to year weather variability and soil input data as causes of uncertainty in regional crop modelling.

Since the studies presented in chapters 3 and 4 considered simple aggregation approaches, i.e. just averaging for weather data and choosing a representative soil profile for spatial soil units, further research should be undertaken in order to gain a deeper insight into the way uncertainty is introduced by other more elaborated aggregation methods such as sampling in geographic space or sampling in probability space (Hansen and Jones, 2000).

The computation of water balance is highly dependent on the method used for calculating evapotranspiration. Therefore, a systematic analysis of the influence of evapotranspiration calculations and their interaction with input data resolution as source of uncertainty for regional yield simulations is highly recommendable. In this respect, and based on preliminary calculations of the difference between different pedotransfer functions used in the study presented in chapter 4, it is recommended to carry out a systematic analysis of the impact of different pedotransfer functions/rules as uncertainty source in soil input data.

The methodology used to generate sub-regional parameter sets in chapter 2, i.e. a brute force search algorithm based on the minimizing of RMSE between simulated and observed yields, is very rudimentary and might not be recommendable to be applied in further regional studies. More sophisticated calibration methodologies like the Bayesian approach (Van Oijen et al., 2005) might offer a more comprehensive insight into the uncertainties related to the parameters which might influence at most the depiction of the spatial variability of crop growth and yield in crop modelling regional applications.

Based on the results of chapter 2 it is recommended to further investigate the effects of considering sub-regional differences in the calibration process not using statistics but well documented field trial data.

In general, it might be recommendable to undertake similar studies considering other (non-cereal) crops or even crop rotations. For this purpose, the scaling of input data referring to management practices might play a very important role. In this respect, it is highly recommendable to undertake studies searching for meaningful strategies to scale input management data.

Finally, although yield is the most important variable for assessment, considering other output variables in further studies (e.g. soil-water dynamics) might offer a better understanding of the dynamics of plant growing at regional level.

5.6 Conclusion

After having systematically addressed the effects of weather and soil input data aggregation, and the choice of model calibration strategy on regional crop model simulation results, the following general conclusions can be drawn:

1. In humid regions, with temperate climate conditions during growing period and homogeneous topography, low resolution climate data seem to be sufficient for climate impact assessment on water limited crop yield.
2. The utilization of various crop models differing in complexity and modelling approaches should become a requirement for every regional impact assessment study since the uncertainties introduced by the model choice have been shown in this study to be more important than the uncertainties caused by the input data resolution.
3. Since the quality of input data as well as data used for model calibration is essential for producing accurate and plausible regional model simulation results, it is indispensable for the crop modelling community to tighten cooperation links with data collectors and data providers in order to obtain data that are suitable for regional crop model applications. For this purpose data collection and data administration protocols should be implemented at regional and global level (e.g. within projects such as MACSUR and AgMIP).
4. Reasonable and useful regional crop modelling work cannot be undertaken by isolated research groups depending on limited resources and hampered by the specific requirements of funding agencies. Multinational and multidisciplinary scientific work focusing on the development of common strategies to tackle issues such as data scarcity and supporting the know-how exchange seems to be a good basis for the generation of knowledge which can be productively used by scientists to provide robust information for decision makers.
5. The use of easy to understand means such as bean plots can support model evaluation through visual assessment and guide a more elaborated quantitative assessment of the effect of input data resolution on the uncertainty of regional yield simulations.
6. The influence of management practices is still not entirely considered in regional simulation yield assessments. Therefore, the search for meaningful strategies to scale management input data for regional modelling applications need to be urgently tackled by the crop modelling community.

6. References

Adam, M., Belhouchette, H., Corbeels, M., Ewert, F., Perrin, A., Casellas, E., Celette, F. and Wery, J., 2012. Protocol to support model selection and evaluation in a modular crop modelling framework: An application for simulating crop response to nitrogen supply. Comput. Electron. Agr., 86(0): 43-54.

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. Modell., 222(1): 131-143.

Addiscott, T. and Whitmore, A., 1991a. Simulation of solute leaching in soils of differing permeabilities. Soil Use Manag., 7(2): 94-102.

Addiscott, T.M. and Whitmore, A.P., 1991b. Simulation of solute leaching in soils of differing permeabilities. Soil Use Manag., 7(2): 94-102.

AG-Boden, 2005. Ad hoc Arbeitsgruppe Boden: Bodenkundliche Kartieranleitung. Bundesanstalt für Geowissenschaften und Rohstoffe in Zusammenarbeit mit den Staatlichen Geologischen Diensten, Hannover, 438 pp.

Ainsworth, E.A. and Long, S.P., 2004. What have we learned from 15 years of free-air CO_2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO_2. New Phytol., 165(2): 351-372.

Allen, R.G., Dirk, R. and Smith, M., 1998a. FAO Irrigation and Drainage Paper No. 56. Rome, Italy.

Allen, R.G., Dirk, R. and Smith, M., 1998b. FAO Irrigation and Drainage Paper No. 56, Rome, Italy, 174 pp.

Andersen, E., Elbersen, B., Hazeu, G. and et al., 2010. The environmental component, the farming systems component and the socio-economic component of the final version of the SEAMLESS database, D4.3.5-D4.4.5-D4.5.4, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org, pp. 401-401.

Angulo, C., Rötter, R., Lock, R., Enders, A., Fronzek, S. and Ewert, F., 2013a. Implication of crop model calibration strategies for assessing regional impacts of climate change in Europe. Agric. For. Meteorol., 170: 32-46.

Angulo, C., Rötter, R., Trnka, M., Pirttioja, N., Gaiser, T., Hlavinka, P. and Ewert, F., 2013b. 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. Nature Clim. Change, 3(9): 827-832.

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

Bakker, M.M., Govers, G., Ewert, F., Rounsevell, M. and Jones, R., 2005. Variability in regional wheat yields as a function of climate, soil and economic variables: Assessing the risk of confounding. Agric. Ecosyst. Environ., 110(3–4): 195-209.

Batchelor, W.D., Basso, B. and Paz, J.O., 2002. Examples of strategies to analyze spatial and temporal yield variability using crop models. Eur. J. Agron., 18(1-2): 141-158.

Bechini, L., Bocchi, S. and Maggiore, T., 2003. Spatial interpolation of soil physical properties for irrigation planning. A simulation study in northern Italy. Eur. J. Agron., 19(1): 1-14.

BK50, 2004. Bodenkarte 1 : 50.000 (BK50). Karte der schutzwürdigen Böden. Bearbeitungsmaßstab 1:50.000. CD-ROM. Geologischer Dienst Nordrhein-Westfalen. Krefeld.

Bondeau, A., Smith, P.C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., Lotze-Campen, H., MüLler, C., Reichstein, M. and Smith, B., 2007. Modelling the role of agriculture for the 20th century global terrestrial carbon balance. Glob. Change Biol., 13(3): 679-706.

Boogaard, H.L., Diepen, C.A., Rotter, R.P., Cabrera, J.M.C.A. and Laar, H.H., 1998. WOFOST 7.1; user's guide for the WOFOST 7.1 crop growth simulation model and WOFOST Control Center 1.5 (No. 52). SC-DLO Wageningen, 144 pp.

Boons-Prins, E.R., de Koning, G.H.J., Diepen, C.A.v. and de Vries, P., 1993. Crop specific simulation parameters for yield forecasting across the European Community. CABO-DLO, no 32., Wageningen, 160 pp.

Brisson, N., Mary, B., Ripoche, D., Jeuffroy, M.H., Ruget, F., Nicoullaud, B., Gate, P., Devienne-Barret, F., Antonioletti, R., Durr, C. and others, 1998. STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn. Agronomie, 18(5-6): 311-346.

Challinor, A.J., Ewert, F., Arnold, S., Simelton, E. and Fraser, E., 2009a. Crops and Climate Change: Progress, Trends, and Challenges in Simulating Impacts and Informing Adaptation. J. Exp. Bot., 60(10): 2775-2789.

Challinor, A.J., Osborne, T., Morse, A., Shaffrey, L., Wheeler, T., Weller, H. and Vidale, P.L., 2009b. Methods and Resources for Climate Impacts Research. Bull. Am. Meteorol. Soc., 90(6): 836-848.

Challinor, A.J., Simelton, E.S., Fraser, E.D.G., Hemming, D. and Collins, M., 2010. Increased crop failure due to climate change: assessing adaptation options using models and socio-economic data for wheat in China. Environ. Res. Lett., 5: 034012-034012.

Challinor, A.J., Wheeler, T.R., Craufurd, P.Q., Slingo, J.M. and Grimes, D.I.F., 2004. Design and optimisation of a large-area process-based model for annual crops. Agric. For. Meteorol., 124(1-2): 99-120.

Challinor, A.J., Wheeler, T.R., Slingo, J.M., Craufurd, P.Q. and Grimes, D.I.F., 2005. Simulation of crop yields using ERA-40: limits to skill and nonstationarity in weather-yield relationships. J. Appl. Meteorol., 44(4): 516-531.

Dark, S.J. and Bram, D., 2007. The modifiable areal unit problem (MAUP) in physical geography. Progress in Physical Geography, 31(5): 471-479.

De Wit, A.J.W., Baruth, B., Boogaard, H.L., van Diepen, K., van Kraalingen, D., Micale, F., te Roller, J., Supit, I. and van den Wijngaart, R., 2010. Using ERA-INTERIM for regional crop yield forecasting in Europe. Clim. Res., 44(1): 41-53.

De Wit, A.J.W., Boogaard, H.L. and Diepen, C.A., 2005. Spatial resolution of precipitation and radiation: The effect on regional crop yield forecasts. Agric. For. Meteorol., 135(1-4): 156-168.

De Wit, C.T. and van Keulen, H., 1987. Modelling production of field crops and its requirements. Geoderma, 40(3–4): 253-265.

Destatis, 2010. Regionaldatenbank Statistische Ämter des Bundes und der Länder.

Donatelli, M., Russell, G., Rizzoli, A.E., Acutis, M., Adam, M., Athanasiadis, I.N., Balderacchi, M., Bechini, L., Belhouchette, H., Bellocchi, G., Bergez, J.-E., Botta, M., Braudeau, E., Bregaglio, S., Carlini, L., Casellas, E., Celette, F., Ceotto, E., Charron-Moirez, M.H., Confalonieri, R., Corbeels, M., Criscuolo, L., Cruz, P., Guardo, A., Ditto, D., Dupraz, C., Duru, M., Fiorani, D., Gentile, A., Ewert, F., Gary, C., Habyarimana, E., Jouany, C., Kansou, K., Knapen, R., Filippi, G.L., Leffelaar, P.A., Manici, L., Martin, G., Martin, P., Meuter, E., Mugueta, N., Mulia, R., Noordwijk, M., Oomen, R., Rosenmund, A., Rossi, V., Salinari, F., Serrano, A., Sorce, A., Vincent, G., Theau, J.-P., Thérond, O., Trevisan, M., Trevisiol, P., Evert, F.K., Wallach, D., Wery, J. and Zerourou, A., 2010. A Component-Based Framework for Simulating Agricultural Production and Externalities. In: F.M. Brouwer and M.K. Ittersum (Editors), Environmental and Agricultural Modeling. Springer Netherlands, Dordrecht, pp. 63-108.

Drebs, A., Nordlund, A., Karlsson, P., Helminen, J. and Rissanen, P., 2002. Climatological Statistics of Finland 1971-2000. Imlastotilatoja Suomesta No. 2002:1, Finnish Meteorological Institute, Helsinki, 99 pp.

DWD, 2010. (Deutscher Wetter Dienst) German Meteorological Service WebWerdis – Weather Request and Distribution System.

Easterling, W.E., Aggarwal, P.K., Punsalmaa, B., Keith, B., Lin, E., Howden, M., Kirilenko, A., Morton, J., Soussana, J.-F., Schmidhuber, J. and Tubiello, F., 2007. Food, fibre and forest products. In: M.L. Parry, O. Canziani, J. Palutikov, P. van der Linden and C. Hanson (Editors), Climate Change 2007: impacts, adaptation and vulnerability : contribution of Working Group II to the fourth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, pp. 273-313.

Easterling, W.E., Mearns, L.O., Hays, C.J. and Marx, D., 2001. Comparison of Agricultural Impacts of Climate Change Calculated from High and Low Resolution Climate Change Scenarios: Part II. Accounting for Adaptation and CO_2 Direct Effects. Clim. Change, 51(2): 173-197.

Easterling, W.E., Weiss, A., Hays, C.J. and Mearns, L.O., 1998. Spatial scales of climate information for simulating wheat and maize productivity: the case of the US Great Plains. Agric. For. Meteorol., 90(1–2): 51-63.

Eurostat, 2010. European Comission Statistics.

Ewert, F., Angulo, C., Rumbaur, C., Lock, R., Enders, A., Adenäuer, M., Heckelei, T., Van Ittersum, M.K., Wolf, J. and Rötter, R., 2011a. Assessing the adaptive capacity of agriculture in the Netherlands to the impacts of climate change under different market and policy scenarios, Dutch Research Program Climate Change and Spatial Planning, 49 pp.

Ewert, F., Keulen, H.v., Ittersum, M.K.v., Giller, K.E., Leffelaar, P.A. and Roetter, R.P., 2006. Multi-scale analysis and modelling of natural resource management options. A. Voinov, A. Jakeman, A. Rizzoli (Eds.), Proceedings of the iEMSs Third Biannual Meeting “Summit on Environmental Modelling and Software”. International Environmental Modelling and Software Society, Burlington.

Ewert, F., Rodriguez, D., Jamieson, P., Semenov, M.A., Mitchell, R.A.C., Goudriaan, J., Porter, J.R., Kimball, B.A., Pinter, P.J., Manderscheid, R., Weigel, H.J., Fangmeier, A., Fereres, E. and Villalobos, F., 2002. Effects of elevated CO_2 and drought on wheat: testing crop simulation models for different experimental and climatic conditions. Agric. Ecosyst. Environ., 93(1-3): 249-266.

Ewert, F., Rounsevell, M.D.A., Reginster, I., Metzger, M.J. and Leemans, R., 2005. Future scenarios of European agricultural land use: I. Estimating changes in crop productivity. Agric. Ecosyst. Environ., 107(2-3): 101-116.

Ewert, F., Van Ittersum, M.K., Heckelei, T., Therond, O., Bezlepkina, I. and Andersen, E., 2011b. Scale changes and model linking methods for integrated assessment of agri-environmental systems. Agric. Ecosyst. Environ., 142(1-2): 6-17.

Ewert, F., Van Oijen, M. and Porter, J.R., 1999. Simulation of growth and development processes of spring wheat in response to CO_2 and ozone for different sites and years in Europe using mechanistic crop simulation models. Eur. J. Agron., 10(3-4): 231-247.

Faivre, R., Leenhardt, D., Voltz, M., Benoît, M., Papy, F., Dedieu, G. and Wallach, D., 2004. Spatialising crop models. Agronomie, 24(4): 205-217.

Farré, I., Van Oijen, M., Leffelaar, P.A. and Faci, J.M., 2000. Analysis of maize growth for different irrigation strategies in northeastern Spain. Eur. J. Agron., 12(3-4): 225-238.

Fischer, G., Shah, M., N. Tubiello, F. and Van Velhuizen, H., 2005. Socio-economic and climate change impacts on agriculture: an integrated assessment, 1990–2080. Philos. Trans. R. Soc. London, Ser. B, 360(1463): 2067-2083.

Folberth, C., Yang, H., Wang, X. and Abbaspour, K.C., 2012. Impact of input data resolution and extent of harvested areas on crop yield estimates in large-scale agricultural modeling for maize in the USA. Ecol. Modell., 235-236: 8-18.

Frère, M. and Popov, G.F., 1979. Agrometeorological crop monitoring and forecasting, FAO Plant Production and Protection Paper (FAO). No. 17, 67 pp.

Gaiser, T., Abdel-Razek, M. and Bakara, H., 2009. Modeling carbon sequestration under zero-tillage at the regional scale. II. The influence of crop rotation and soil type. Ecol. Modell., 220(23): 3372-3379.

Gaiser, T., Perkons, U., Küpper, P.M., Kautz, T., Uteau-Puschmann, D., Ewert, F., Enders, A. and Krauss, G., 2013. Modeling biopore effects on root growth and biomass production on soils with pronounced sub-soil clay accumulation. Ecol. Modell., 256: 6-15.

Garcia, R., Kanemasu, E.T., Blad, B.L., Bauer, A., Hatfield, J.L., Major, D.J., Reginato, R.J. and Hubbard, K.G., 1988. Interception and use efficiency of light in winter wheat under different nitrogen regimes. Agric. For. Meteorol., 44(2): 175-186.

Gassman, P.W., Williams, J.R., Benson, V.W., Izaurralde, R.C., Hauck, L.M., Allan Jones, C., Atwood, J.D., Kiniry, J.R. and Flowers, J.D., 2004. Historical development and applications of the EPIC and APEX models, ASAE/CSAE Meeting. Paper No. 042097. ASAE, St. Joseph.

Goyne, P.J. and Hare, 1993. Radiation interception, radiation use efficiency and growth of barley cultivars. Aust. J. Agric. Res., 44(6): 1351-1366.

Hakala, K., Jauhiainen, L., Himanen, S.J., Rötter, R., Salo, T. and Kahiluoto, H., 2011. Sensitivity of barley varieties to weather in Finland. J. Agric. Sci., 150(02): 145-160.

Hansen, J.W. and Jones, J.W., 2000. Scaling-up crop models for climate variability applications. Agric. Syst., 65(1): 43-72.

Hansen, S., Abrahamsen, P., Petersen, C.T. and Styczen, M., 2012. DAISY: model use, calibration and validation. Trans. ASABE, 55(4): 1315-1333.

Harrison, P.A. and Butterfield, R.E., 1996. Effects of climate change on Europe-wide winter wheat and sunflower productivity. Clim. Res., 7(1996): 225-241.

Hartwich, R., Behrens, J., Eckelmann, W., Haase, G., Richter, A. and Roeschmann, G., 1995. Bodenübersichtskarte der Bundesrepublik Deutschland 1: 1 000 000. Erläuterungen, Textlegende und Leitprofile. Bundesanstalt für Geowissenschaften und Rohstoffe, Hannover.

Hazeu, G.W., Elbersen, B., Andersen, E., Baruth, B., Van Diepen, K. and Metzger, M.J., 2010. A biophysical typology for a spatially-explicit agri-environmental modeling framework. In: F. Brouwer and M.K. Van Ittersum (Editors), Environmental and agricultural modelling: integrated approaches for policy impact assessment. Springer, Dordrecht, pp. 159-187.

Hermans, C.M.L., Geijzendorffer, I.R., Ewert, F., Metzger, M.J., Vereijken, P.H., Woltjer, G.B. and Verhagen, A., 2010. Exploring the future of European crop production in a liberalised market, with specific consideration of climate change and the regional competitiveness. Ecol. Modell., 221(18): 2177-2187.

Heuvelink, G.B.M. and Webster, R., 2001. Modelling soil variation: past, present, and future. Geoderma, 100(2001): 269-301.

Hijmans, R.J., 2003. The effect of climate change on global potato production. American Journal of Potato Research, 80(4): 271-279.

Holt, D., Steel, D.G., Tranmer, M. and Wrigley, N., 1996. Aggregation and Ecological Effects in Geographically Based Data. Geographical Analysis, 28(3): 244-261.

Hui, C., 2009. A Bayesian Solution to the Modifiable Areal Unit Problem. In: A.-E. Hassanien, A. Abraham and F. Herrera (Editors), Foundations of Computational Intelligence. Studies in Computational Intelligence. Springer Berlin Heidelberg, pp. 175-196.

Iglesias, A., Quiroga, S. and Schlickenrieder, J., 2010. Climate change and agricultural adaptation: assessing management uncertainty for four crop types in Spain. Clim. Res., 44(1): 83-94.

Izaurralde, R.C., Williams, J.R., McGill, W.B., Rosenberg, N.J. and Jakas, M.C., 2006. Simulating soil C dynamics with EPIC: Model description and testing against long-term data. Ecol. Modell., 192(3-4): 362-384.

Jaggard, K.W., Allison, M.F., Clark, C.J.A., Todd, A.D. and Smith, H.G., 2003. The effect of nitrogen supply and virus yellows infection on the growth, yield and processing quality of sugarbeet (Beta vulgaris). J. Agric. Sci., 139(02).

Jagtap, S.S. and Jones, J.W., 2002. Adaptation and evaluation of the CROPGRO-soybean model to predict regional yield and production. Agric. Ecosyst. Environ., 93(1-3): 73-85.

Jamieson, P.D., Porter, J.R., Goudriaan, J., Ritchie, J.T., van Keulen, H. and Stol, W., 1998. A comparison of the models AFRCWHEAT2, CERES-Wheat, Sirius, SUCROS2 and SWHEAT with measurements from wheat grown under drought. Field Crops Res., 55(1–2): 23-44.

Janssen, S., Andersen, E., Athanasiadis, I.N. and Van Ittersum, M.K., 2009. A database for integrated assessment of European agricultural systems. Environmental Science & Policy, 12(5): 573-587.

Johansson, B. and Chen, D., 2003. The influence of wind and topography on precipitation distribution in Sweden: statistical analysis and modelling. Int. J. Climatol., 23(12): 1523-1535.

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J. and Ritchie, J.T., 2003a. The DSSAT cropping system model. Eur. J. Agron., 18(3-4): 235-265.

Jones, P.D., Lister, D.H., Jaggard, K.W. and Pidgeon, J.D., 2003b. Future climate impact on the productivity of sugar beet (Beta vulgaris L.) in Europe. Clim. Change, 58(1): 93-108.

JRC, 1998. Estimation of the phenological calendar, Kc-curve and temperature sums for cereals, sugar beet, potato, sunflower and rapeseed across Pan Europe, Turkey and the Maghreb countries by means of transfer procedures, Willekens, A., van Orshoven, J., Feyen, J. (Eds.), European Commission (JRC-SAI) Agrometeorological transfer procedures: Vol 1-3.

Kampstra, P., 2008. Beanplot: A boxplot alternative for visual comparison of distributions. J. Stat. Softw., 28(2008): 1-9, (code snippet 1).

Kruijt, B., Witte, J.-P.M., Jacobs, C.M.J. and Kroon, T., 2008. Effects of rising atmospheric CO_2 on evapotranspiration and soil moisture: A practical approach for the Netherlands. J. Hydrol., 349(3-4): 257-267.

Langensiepen, M., Hanus, H., Schoop, P. and Gräsle, W., 2008. Validating CERES-wheat under North-German environmental conditions. Agric. Syst., 97(1-2): 34-47.

Launay, M. and Guerif, M., 2005. Assimilating remote sensing data into a crop model to improve predictive performance for spatial applications. Agric. Ecosyst. Environ., 111(1-4): 321-339.

Lawless, C., Semenov, M.A. and Jamieson, P.D., 2008. Quantifying the effect of uncertainty in soil moisture characteristics on plant growth using a crop simulation model. Field Crops Res., 106(2): 138-147.

Leakey, A.D.B., Ainsworth, E.A., Bernacchi, C.J., Rogers, A., Long, S.P. and Ort, D.R., 2009. Elevated CO_2 effects on plant carbon, nitrogen, and water relations: six important lessons from FACE. J. Exp. Bot., 60(10): 2859-2876.

Leenhardt, D., Voltz, M., Bornand, M. and Webster, R., 1994. Evaluating soil maps for prediction of soil water properties. European J. Soil Sci., 45 (3): 293-301.

Leenhardt, D., Wallach, D., Le Moigne, P., Guérif, M., Bruand, A. and Casterad, M.A., 2006. Using crop models for multiple fields, Working with Dynamic Crop Models. Evaluation, Analysis, Parameterization and Application. Elsevier, pp. 209-248.

Lehuger, S., Gabrielle, B., Oijen, M., Makowski, D., Germon, J.C., Morvan, T. and Hénault, C., 2009. Bayesian calibration of the nitrous oxide emission module of an agro-ecosystem model. Agric. Ecosyst. Environ., 133(3-4): 208-222.

Lobell, D.B., Burke, M.B., Tebaldi, C., Mastrandrea, M.D., Falcon, W.P. and Naylor, R.L., 2008. Prioritizing Climate Change Adaptation Needs for Food Security in 2030. Science, 319(5863): 607-610.

Lobell, D.B., Cassman, K.G. and Field, C.B., 2009. Crop Yield Gaps: Their Importance, Magnitudes, and Causes. Annual Review of Environment and Resources, 34(1): 179-204.

Long, S.P., 2006. Food for Thought: Lower-Than-Expected Crop Yield Stimulation with Rising CO_2 Concentrations. Science, 312(5782): 1918-1921.

LSV, 2012. Landessortenversuche Winterweizen Erntejahr 2003 – 2005. Landwirtschaftskammer NRW, Bonn.

Makkink, G.F., 1957. Testing the Penman formula by means of lysimeters. J. Inst. Water Eng, 11(3): 277-288.

Manderscheid, R. and Weigel, H.-J., 2007. Drought stress effects on wheat are mitigated by atmospheric CO_2 enrichment. Agronomy for Sustainable Development, 27(2): 79-87.

Mattila, P., Rankinen, K., Grönroos, J., Siimes, K., Karhu, E., Laitinen, K., Granlund, K., Ekholm, P. and Antikainen, R., 2007. Viljelytoimenpiteet ja vesistökuormitus ympäristötukitiloilla v. 2003-2005, Suomen ympäristö 40/2007, 101 pp.

McMaster, G.S., White, J.W., Hunt, L.A., Jamieson, P.D., Dhillon, S.S. and Ortiz-Monasterio, J.I., 2008. Simulating the Influence of Vernalization, Photoperiod and Optimum Temperature on Wheat Developmental Rates. Annals of Botany, 102(4): 561-569.

Mearns, L.O., Easterling, W., Hays, C. and Marx, D., 2001. Comparison of Agricultural Impacts of Climate Change Calculated from High and Low Resolution Climate Change Scenarios: Part I. The Uncertainty Due to Spatial Scale. Clim. Change, 51(2): 131-172.

Mearns, L.O., Giorgi, F., McDaniel, L. and Shields, C., 2003. Climate scenarios for the southeastern US based on GCM and regional model simulations. Clim. Change, 60(1): 7-35.

Mearns, L.O., Mavromatis, T., Tsvetsinskaya, E., Hays, C. and Easterling, W., 1999. Comparative responses of EPIC and CERES crop models to high and low spatial resolution climate change scenarios. J. Geophys. Res., 104(D6): 6623-6646.

Meehl, A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J.F.B., Stouffer, R.J., Karl and Taylor, E., 2007. The WCRP CMIP3 multimodel dataset: A new ero in climate change research.

Metzger, M.J., Bunce, R.G.H., Jongman, R.H.G., Mücher, C.A. and Watkins, J.W., 2005. A climatic stratification of the environment of Europe. Global Ecology and Biogeography, 14(6): 549-563.

Mignolet, C., Schott, C. and Benoît, M., 2004. Spatial dynamics of agricultural practices on a basin territory: a retrospective study to implement models simulating nitrate flow. The case of the Seine basin. Agronomie, 24(4): 219-236.

Moen, T.N., Kaiser, H.M. and Riha, S.J., 1994. Regional yield estimation using a crop simulation model: Concepts, methods, and validation. Agric. Syst., 46(1): 79-92.

Monteith, J.L. and Greenwood, D.J., 1986. How Do Crops Manipulate Water Supply and Demand? Philos. Trans. R. Soc. London, Ser. A, 316(1537): 245-259.

Monteith, J.L. and Moss, C.J., 1977. Climate and the Efficiency of Crop Production in Britain. Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 281(980): 277-294.

Nakicenovic, N. and Swart, R., Ed., 2000. Special Report on Emissions Scenarios, Cambridge University Press, 612 pp.

Nendel, C., Wieland, R., Mirschel, W., Specka, X., Guddat, C. and Kersebaum, K.C., 2013. Simulating regional winter wheat yields using input data of different spatial resolution. Field Crops Res., 145((2013)): 67-77.

Niu, X., Easterling, W., Hays, C.J., Jacobs, A. and Mearns, L., 2009. Reliability and input-data induced uncertainty of the EPIC model to estimate climate change impact on sorghum yields in the U.S. Great Plains. Agric. Ecosyst. Environ., 129(1-3): 268-276.

O'neill, R.V. and King, A.W., 1998. Homage to St. Michael; or, why are there so many books on scale? In: D.L. Peterson and V.T. Parker (Editors), Ecological Scale: Theory and Applications. Complexity in ecological systems. . Columbia University Press, New York, pp. 3-15.

Olesen, J.E. and Bindi, M., 2002. Consequences of climate change for European agricultural productivity, land use and policy. Eur. J. Agron., 16(4): 239-262.

Olesen, J.E., Bocher, P.K. and Jensen, T., 2000. Comparison of scales of climate and soil data for aggregating simulated yields of winter wheat in Denmark. Agric. Ecosyst. Environ., 82(1-3): 213-228.

Olesen, J.E., Carter, T.R., Díaz-Ambrona, C.H., Fronzek, S., Heidmann, T., Hickler, T., Holt, T., Minguez, M.I., Morales, P., Palutikof, J.P., Quemada, M., Ruiz-Ramos, M., Rubæk, G.H., Sau, F., Smith, B. and Sykes, M.T., 2007. Uncertainties in projected impacts of climate change on European agriculture and terrestrial ecosystems based on scenarios from regional climate models. Clim. Change, 81(Suppl. 1): 123-143.

Palosuo, T., Kersebaum, K.C., Angulo, C., Hlavinka, P., Moriondo, M., Olesen, J.E., Patil, R.H., Ruget, F., Rumbaur, C., Takáč, J., Trnka, M., Bindi, M., Çaldağ, B., Ewert, F., Ferrise, R., Mirschel, W., Şaylan, L., Šiška, 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(3): 103-114.

Palva, R., Rankinen, K., Granlund, K., Grönroos, J., Nikander, A. and Rekolainen, S., 2001. Maatalouden ympäristötuen toimenpiteiden toteutuminen ja vaikutukset vesistökuormitukseen vuosina 1995–1999, Suomen ympäristö 478, 92 pp.

Parry, M., Rosenzweig, C., Iglesias, A., Fischer, G. and Livermore, M., 1999. Climate change and world food security: a new assessment. Global Environ. Change, 9, Supplement 1: S51-S67.

Parry, M.L., Rosenzweig, C., Iglesias, A., Livermore, M. and Fischer, G., 2004. Effects of climate change on global food production under SRES emissions and socio-economic scenarios. Global Environ. Change, 14(1): 53-67.

Penman, H.L., 1956. Evaporation-an introductory survey. Netherlands Journal of Agricultural Science(4): 9-29.

Penning de Vries, F.W.T., Jansen, D.M., Ten Berge, H.F.M. and Bakema, A., 1989. Simulation of ecophysiological processes of growth in several annual crops, Simulation Monographs 29. Pudoc, Wageningen.

Porter, J.R. and Gawith, M., 1999. Temperatures and the growth and development of wheat: a review. Eur. J. Agron., 10(1): 23-36.

Porter, J.R., Jamieson, P.D. and Wilson, D.R., 1993. Comparison of the wheat simulation models Afrcwheat2, Ceres-wheat and Swheat for non-limiting conditions of crop growth. Field Crops Res., 33(1–2): 131-157.

Porter, J.R. and Semenov, M.A., 2005. Crop responses to climatic variation. Philos. Trans. R. Soc. London, Ser. B, 360(1463): 2021-2035.

Priestley, C.H.B. and Taylor, R.J., 1972. On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. Monthly Weather Rev., 100(2): 81-92.

Pyykkönen, S., Grönroos, J., Rankinen, K., Laitinen, P., Karhu, E. and Granlund, K., 2004. Ympäristötuen mukaiset viljelymenetelmät ja niiden vaikutukset vesistökuormitukseen vuosina 2000-2002, 119 pp.

Reidsma, P., Ewert, F., Boogaard, H. and Diepen, K., 2009a. Regional crop modelling in Europe: The impact of climatic conditions and farm characteristics on maize yields. Agric. Syst., 100(1-3): 51-60.

Reidsma, P., Ewert, F., Oude Lansink, A. and Leemans, R., 2009b. Vulnerability and adaptation of European farmers: a multi-level analysis of yield and income responses to climate variability. Regional Environmental Change, 9(1): 25-40.

Reinds, G.J., Van Oijen, M., Heuvelink, G. and Kros, H., 2008. Bayesian calibration of the VSD soil acidification model using European forest monitoring data. Geoderma, 146(3-4): 475-488.

Reyenga, P.J., Howden, S.M., Meinke, H. and McKeon, G.M., 1999. Modelling global change impacts on wheat cropping in south-east Queensland, Australia. Environ. Model. Software, 14(4): 297-306.

Rodriguez, D., Ewert, F., Goudriaan, J., Manderscheid, R., Burkart, S. and Weigel, H.J., 2001. Modelling the response of wheat canopy assimilation to atmospheric CO_2 concentrations. New Phytol., 150(2): 337-346.

Rosenthal, W.D., Hammer, G.L. and Butler, D., 1998. Predicting regional grain sorghum production in Australia using spatial data and crop simulation modelling. Agric. For. Meteorol., 91(3-4): 263-274.

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. For. Meteorol., 170(2013): 166-182.

Rosenzweig, C. and Parry, M.L., 1994. Potential impact of climate change on world food supply. Nature, 367(6459): 133-138.

Rossberg, D., Michel, V., Graf, R. and Neukampf, R., 2007. Definition von Boden-Klima-Räumen für die Bundesrepublik Deutschland. Nachrichtenblatt des Deutschen Pflanzenschutzdienstes, 59(7): 155-161.

Rötter, R., Ewert, F., Palosuo, T., Bindi, M., Kersebaum, K.C., Olesen, J.E., Trnka, M., Van Ittersum, M.K., Janssen, S., Rivington, M., Semenov, M., Wallach, D., Porter, J.R., Stewart, D., Verhagen, J., Angulo, C., Gaiser, T., Nendel, C., Martre, P. and De Wit, A., 2013a. Challenges for Agro-Ecosystem Modelling in Climate Change Risk Assessment for major European Crops and Farming systems. Proceedings of the Impacts World 2013 conference at Potsdam, Germany, May 2013, 555-564. DOI: 10.2312/pik.2013.001.

Rötter, R., Palosuo, T., Kersebaum, K.C., Angulo, C., Bindi, M., Ewert, F., Ferrise, R., Hlavinka, P., Moriondo, M., Nendel, C., Olesen, J.E., Patil, R.H., Ruget, F., Takáč, J. and Trnka, M., 2012a. Simulation of spring barley yield in different climatic zones of Northern and Central Europe. A comparison of nine crop models. Field Crops Res., 133(2012): 23-36.

Rötter, R.P., Carter, T.R., Olesen, J.E. and Porter, J.R., 2011a. Crop-climate models need an overhaul. Nature Clim. Change, 1(4): 175-177.

Rötter, R.P., Höhn, J.G. and Fronzek, S., 2013b. Projections of climate change impacts on crop production: A global and a Nordic perspective. Acta Agriculturae Scand., Section A, 62(4): 166-180.

Rötter, R.P., Palosuo, T., Kersebaum, K.C., Angulo, C., Bindi, M., Ewert, F., Ferrise, R., Hlavinka, P., Moriondo, M., Nendel, C., Olesen, J.E., Patil, R.H., Ruget, F., Takáč, J. and Trnka, M., 2012b. Simulation of spring barley yield in different climatic zones of Northern and Central Europe. A comparison of nine crop models. Field Crops Res., 133(2012): 23-36.

Rötter, R.P., Palosuo, T., Pirttioja, N.K., Dubrovsky, M., Salo, T., Fronzek, S., Aikasalo, R., Trnka, M., Ristolainen, A. and Carter, T.R., 2011b. What would happen to barley production in Finland if global warming exceeded 4° C? A model-based assessment. Eur. J. Agron., 35(2011): 205-214.

Rötter, R.P. and Van De Geijn, S.C., 1999. Climate Change Effects on Plant Growth, Crop Yield and Livestock. Clim. Change, 43(4): 651-681.

Rötter, R.P. and Van Diepen, C.A., 1994. Rhine Basin Study: Land Use Projections Based on Biophysical and Socioeconomic Analysis. Climate change impact on crop yield potentials and water use, 152 pp.

Salo, T. and et al., in preparation. Comparing the performance of eleven agro-ecosystems models in predicting crop yield response to nitrogen under Finnish weather conditions.

Savary, S., Teng, P.S., Willocquet, L. and Nutter, F.W., 2006. Quantification and Modeling of Crop Losses: A Review of Purposes. Annu. Rev. Phytopatol., 44(1): 89-112.

Semenov, M.A. and Halford, N.G., 2009. Identifying Target Traits and Molecular Mechanisms for Wheat Breeding Under a Changing Climate. J. Exp. Bot., 60(10): 2791-2804.

Semenov, M.A. and Pilkington-Bennett, S., 2012. Validation of ELPIS baseline scenarios using ECA&D observed data. In: EGU General Assembly 2012: 22-27 April, 2012 Vienna, Austria; pp 2396.

Semenov, M.A., PilkingtonBennett, S. and Calanca, P., 2013. Validation of ELPIS 1980-2010 baseline scenarios using the observed European Climate Assessment data set. Clim. Res., 57(1): 1-9.

Semenov, M.A. and Shewry, P.R., 2011. Modelling predicts that heat stress, not drought, will increase vulnerability of wheat in Europe. Sci. Rep. 1, 66(2011).

Sheather, S.J., 2004. Density Estimation. Statist. Sci., 19(4): 588-597.

Sheather, S.J. and Jones, M.C., 1991. A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation. Journal of the Royal Statistical Society. Series B (Methodological), 53(3): 683-690.

Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Chapman and Hall, New York.

Slafer, G.A. and Rawson, H.M., 1996. Responses to photoperiod change with phenophase and temperature during wheat development. Field Crops Res., 46(1-3): 1-13.

Spitters, C.J.T., 1990. Crop growth models: their usefulness and limitations. Acta Hortic., 267(1990): 349-368.

Spitters, C.J.T. and Schapendonk, A., 1990. Evaluation of breeding strategies for drought tolerance in potato by means of crop growth simulation. Plant Soil, 123(2): 193-203.

Stockle, C.O., Donatelli, M. and Nelson, R., 2003. CropSyst, a cropping systems simulation model. Eur. J. Agron., 18(2003): 289-307.

Stockle, C.O., Williams, J.R., Rosenberg, N.J. and Jones, C.A., 1992. A method for estimating the direct and climatic effects of rising atmospheric carbon dioxide on growth and yield of

crops: Part I-Modification of the EPIC model for climate change analysis. Agric. Syst., 38(3): 225-238.
Tao, F., Yokozawa, M. and Zhang, Z., 2009. Modelling the impacts of weather and climate variability on crop productivity over a large area: A new process-based model development, optimization, and uncertainties analysis. Agric. For. Meteorol., 149(5): 831-850.
Teixeira, E.I., Fischer, G., van Velthuizen, H., Walter, C. and Ewert, F., 2013. Global hot-spots of heat stress on agricultural crops due to climate change. Agric. For. Meteorol., 170(2013): 206-215.
Therond, O., Hengsdijk, H., Casellas, E., Wallach, D., Adam, M., Belhouchette, H., Oomen, R., Russell, G., Ewert, F., Bergez, J.-E., Janssen, S., Wery, J. and Van Ittersum, M.K., 2011. Using a cropping system model at regional scale: Low-data approaches for crop management information and model calibration. Agric. Ecosyst. Environ., 142(1-2): 85-94.
Tremblay, M. and Wallach, D., 2004. Comparison of parameter estimation methods for crop models. Agronomie, 24(6-7): 351-365.
Trnka, M., Eitzinger, J., Kapler, P., Dubrovskỳ, M., Semerádová, D., Žalud, Z. and Formayer, H., 2007. Effect of estimated daily global solar radiation data on the results of crop growth models. Sensors, 7(10): 2330-2362.
Trnka, M., Olesen, J.E., Kersebaum, K.C., Skjelvåg, A.O., Eitzinger, J., Seguin, B., Peltonen-Sainio, P., RöTter, R., Iglesias, A., Orlandini, S., Dubrovský, M., Hlavinka, P., Balek, J., Eckersten, H., Cloppet, E., Calanca, P., Gobin, A., VučEtić, V., Nejedlik, P., Kumar, S., Lalic, B., Mestre, A., Rossi, F., Kozyra, J., Alexandrov, V., SemeráDová, D. and žAlud, Z., 2011. Agroclimatic conditions in Europe under climate change. Glob. Change Biol., 17(7): 2298-2318.
Tubiello, F.N., Amthor, J.S., Boote, K.J., Donatelli, M., Easterling, W., Fischer, G., Gifford, R.M., Howden, M., Reilly, J. and Rosenzweig, C., 2007. Crop response to elevated CO_2 and world food supply: A comment on "Food for Thought..." by Long et al., Science 312:1918-1921, 2006. Eur. J. Agron., 26(3): 215-223.
Tubiello, F.N. and Ewert, F., 2002. Simulating the effects of elevated CO_2 on crops: approaches and applications for climate change. Eur. J. Agron., 18(1-2): 57-74.
Tubiello, F.N. and Fischer, G., 2007. Reducing climate change impacts on agriculture: Global and regional effects of mitigation, 2000-2080. Technological Forecasting and Social Change, 74(7): 1030-1056.
Turtola, E. and Lemola, R., 2006. Follow-up of the effectiveness of the agro-environmental programme in Finland. Results of sub-projects 2000-2003. Maa ja elintarivketalous., 59 pp.
Unwin, D.J., 1996. GIS, spatial analysis and spatial statistics. Progress in Human Geography, 20(4): 540-551.
Van Bussel, L.G.J., Ewert, F. and Leffelaar, P.A., 2011a. Effects of data aggregation on simulations of crop phenology. Agric. Ecosyst. Environ., 142(1-2): 75-84.
Van Bussel, L.G.J., Muller, C., Van Keulen, H., Ewert, F. and Leffelaar, P.A., 2011b. The effect of temporal aggregation of weather input data on crop growth models' results. Agric. For. Meteorol., 151(5): 607-619.
Van Der Velde, M., Bouraoui, F. and Aloe, A., 2009. Pan-European regional-scale modelling of water and N efficiencies of rapeseed cultivation for biodiesel production. Glob. Change Biol., 15(1): 24-37.
Van Diepen, C.A., Wolf, J., van Keulen, H. and Rappoldt, C., 1989. WOFOST: a simulation model of crop production. Soil Use Manag., 5(1): 16-24.
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(1-3): 150-165.
Van Ittersum, M.K., Howden, S.M. and Asseng, S., 2003a. Sensitivity of productivity and deep drainage of wheat cropping systems in a Mediterranean environment to changes in CO_2, temperature and precipitation. Agric. Ecosyst. Environ., 97(1-3): 255-273.

Van Ittersum, M.K., Leffelaar, P.A., van Keulen, H., Kropff, M.J., Bastiaans, L. and Goudriaan, J., 2003b. On approaches and applications of the Wageningen crop models. Eur. J. Agron., 18(3-4): 201-234.

Van Oijen, M. and Ewert, F., 1999. The effects of climatic variation in Europe on the yield response of spring wheat cv. Minaret to elevated CO_2 and O_3: an analysis of open-top chamber experiments by means of two crop growth simulation models. Eur. J. Agron., 10(3-4): 249-264.

Van Oijen, M., Rougier, J. and Smith, R., 2005. Bayesian calibration of process-based forest models: bridging the gap between models and data. Tree Physiology, 25(7): 915-915.

Venäläinen, A., Tuomenvirta, H., Pirinen, P. and Drebs, A., 2005. A Basic Finnish climate dataset 1961-2000-description and illustrations. Finnish Meteorological Institute, Reports, 2005: 5-27.

Walker, W.E., Harremoës, P., Rotmans, J., van der Sluijs, J.P., van Asselt, M.B.A., Janssen, P. and Krayer von Krauss, M.P., 2003. Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. Integrated Assessment, 4(1): 5-17.

Wallach, D., Buis, S., Lecharpentier, P., Bourges, J., Clastre, P., Launay, M., Bergez, J.E., Guerif, M., Soudais, J. and Justes, E., 2010. A package of parameter estimation methods and implementation for the STICS crop-soil model. Environ. Model. Software, 26(2011): 386-394.

Wassenaar, T., Lagacherie, P., Legros, J.P. and Rounsevell, M.D.A., 1999. Modelling wheat yield responses to soil and climate variability at the regional scale. Clim. Res., 11(1999): 209-220.

White, J.W., Hoogenboom, G., Kimball, B.A. and Wall, G.W., 2011. Methodologies for simulating impacts of climate change on crop production. Field Crops Res., 124(3): 357-368.

Williams, J.R., Renard, K.G. and Dyke, P.T., 1983. EPIC: A new method for assessing erosion's effect on soil productivity. J. Soil Water Conserv., 38(5): 381-383.

Willmott, C.J., 1981. On the Validation of Models. Phys. Geogr., 2(2): 184-194.

Willmott, C.J., Ackleson, S.G., Davis, R.E., Feddema, J.J., Klink, K.M., Legates, D.R., O'Donnell, J. and Rowe, C.M., 1985. Statistics for the evaluation and comparison of models. J. Geophys. Res., 90(C5): 8995-9005.

Wolf, J., Evans, L.G., Semenov, M.A., Ecker, H. and Iglesias, A., 1996. Comparison of wheat simulation models under climate change. I. Model calibration and sensitivity analyses. Clim. Res., 7(3): 253-270.

Wolf, J. and Oijen, M.v., 2002. Modelling the dependence of European potato yields on changes in climate and CO_2. Agric. For. Meteorol., 112(3-4): 217-231.

Wösten, J.H.M., Lilly, A., Nemes, A. and Le Bas, C., 1999. Development and use of a database of hydraulic properties of European soils. Geoderma, 90(1999): 169-185.

Xiong, W., Holman, I., Conway, D., Lin, E. and Li, Y., 2008. A crop model cross calibration for use in regional climate impacts studies. Ecol. Modell., 213(3-4): 365-380.

Yin, X. and Van Laar, H.H., 2005. Crop Systems Dynamics: An Ecophysiological Simulation Model for Genotype-by-environment Interactions. Wageningen Academic Publishers, Wageningen, 155 pp.

7. Appendixes

Appendix 1. Tabular pedotransfer rules developed for German soils (adapted from _ENREF_4AG-Boden, 2005) (Min=minimum content, Max=maximum content, SAT=water content at saturation (pF=4.2), FC=water content at field capacity (pF=2.5), WP=water content at wilting point (pF=4.2))

Symbol	German classification name	Translation	Clay_Min	Clay_Max	Silt_Min	Silt_Max	Sand_Min	Sand_Max	SAT	FC	WP
Ls2	schwach sandiger Lehm	sandy loam[1]	17	25	40	50	25	43	42	25	18
Ls3	mittel sandiger Lehm	sandy loam[2]	17	25	30	40	35	53	41	24	16
Ls4	stark sandiger Lehm	sandy loam[3]	17	25	15	30	45	68	42	23	16
Lt2	schwach toniger Lehm	clay loam[1]	25	35	30	50	15	45	42	31	22
Lt3	mittel toniger Lehm	clay loam[2]	35	45	30	50	5	35	43	34	26
Lts	sandig-toniger Lehm	sandy clay loam	25	45	15	30	25	60	42	31	22
Lu	schluffiger Lehm	silty loam	17	30	50	65	5	33	42	29	18
Sl2	schwach lehmiger Sand	loamy sand[1]	5	8	10	25	67	85	41	17	7
Sl3	mittel lehmiger Sand	loamy sand[2]	8	12	10	40	48	82	41	20	9
Sl4	stark lehmiger Sand	loamy sand[3]	12	17	10	40	43	78	41	23	12
Slu	schluffig-lehmiger Sand	silty loamy sand	8	17	40	50	33	52	42	26	12
Ss	reiner Sand	sand	0	5	0	10	85	100	42	12	4
St2	schwach toniger Sand	clay sand[1]	5	17	0	10	73	95	40	14	6
St3	mittel toniger Sand	clay sand[2]	17	25	0	15	60	83	42	24	15
Su2	schwach schluffiger Sand	silty sand[1]	0	5	10	25	70	90	41	13	4
Su3	mittel schluffiger Sand	silty sand[2]	0	8	25	40	52	75	42	20	8
Su4	stark schluffiger Sand	silty sand[3]	0	8	40	50	42	60	42	23	9
Tl	lehmiger Ton	loamy clay	45	65	15	30	5	40	44	37	27
Ts2	schwach sandiger Ton	sandy clay[1]	45	65	0	15	20	55	42	37	25
Ts3	mittel sandiger Ton	sandy clay[2]	35	45	0	15	40	65	42	37	23
Ts4	stark sandiger Ton	sandy clay[3]	25	35	0	15	50	75	41	30	19
Tt	reiner Ton	clay	65	100	0	35	0	35	44	39	28
Tu2	schwach schluffiger Ton	silty clay[1]	45	65	30	55	0	25	44	39	29
Tu3	mittel schluffiger Ton	silty clay[2]	30	45	50	65	0	20	43	35	25
Tu4	stark schluffiger Ton	silty clay[3]	25	35	65	75	0	10	42	33	20
Uls	sandig-lehmiger Schluff	sand loamy silt	8	17	50	65	18	42	42	30	13
Us	sandiger Schluff	sandy silt	0	8	50	80	12	50	42	29	10
Ut2	schwach toniger Schluff	clay silt[1]	8	12	65	92	0	27	42	31	12
Ut3	mittel toniger Schluff	clay silt[2]	12	17	65	88	0	23	42	32	13
Ut4	stark toniger Schluff	clay silt[3]	17	25	65	83	0	18	43	33	16
Uu	reiner Schluff	silt	0	8	80	100	0	20	43	32	12

[1]Low content of secondary textural component; [2]Mid content of secondary textural component; [3]High content of secondary textural component