Evaluating the precision of eight spatial sampling schemes in estimating regional mean yields for two crops

*Zhao, Gang1\*;Hoffmann, Holger1;* *Yeluripati, Jagadeesh2,5; Xenia, Specka 3;* *Nendel, Claas4; Coucheney, Elsa 5; Kuhnert,* *Matthias6; Tao,* *Fulu7;* *Constantin, Julie8;* *Raynal, Helene8; Teixeira, Edmar9;* *Grosz, Balázs10; Doro, Luca11;* *Kiese, Ralf12;* *Eckersten, Henrik13; Haas,* *Edwin12;* *Cammarano, Davide14;* *Kassie, Belay14; Moriondo, Marco15; Trombi, Giacomo16**; Bindi, Marco16;* *Biernath, Christian17;* *Heinlein, Florian17; Klein, Christian17;* *Priesack, Eckart17; Lewan, Elisabet5;* *Kersebaum, Kurt-Christian4; Rötter,* *Reimund7; Roggero, Pier Paolo11;* *Wallach, Daniel8; Asseng, Senthold14; Siebert, Stefan1; Gaiser, Thomas1; Ewert, Frank1*

1Crop Science Group, Institute of Crop Science and Resource Conservation (INRES), University of Bonn, Katzenburgweg 5, 53115 Bonn, DE

2The James Hutton Institute, Craigiebuckler, Aberdeen, Scotland AB15 8QH, UK

3Institute of Landscape Biogeochemistry, Leibniz Centre for Agricultural Landscape Research, 15374 Müncheberg, DE

4Institute of Landscape Systems Analysis, Leibniz Centre for Agricultural Landscape Research, 15374 Müncheberg, DE

5Biogeophysics and water quality, Department of Soil and Environment, Swedish University of Agricultural Sciences, Lennart Hjelms väg 9, 750 07 Uppsala, SE

6Institute of Biological and Environmental Sciences, School of Biological Sciences, University of Aberdeen, 23 St Machar Drive, Aberdeen AB24 3 UU, Scotland, UK

7Environmental Impacts Group, Natural Resources Institute Finland (Luke), Jokiniemenkuja1, FI-01301 Vantaa, FI

8INRA, UMR 1248 AGIR & UR0875 MIA-T, F-31326 Auzeville, FR

9Systems Modelling Team (Sustainable Production Group), The New Zealand Institute for Plant and Food Research Limited, Canterbury Agriculture & Science Centre, Gerald St, Lincoln 7608, NZ

10Thünen-Institute of Climate-Smart-Agriculture, Bundesallee 50, 38116 Braunschweig, DE

11Nucleo di Ricerca sulla Desertificazione (NRD) and Dipartimento di Agraria, Università degli Studi di Sassari, Viale Italia 39, 07100 Sassari, IT

12Institute of Meteorology and Climate Research – Atmospheric Environmental Research, Karlsruhe Institute of Technology, Kreuzeckbahnstraße 19, 82467 Garmisch-Partenkirchen, DE

13Department of Crop Production Ecology, Swedish University of Agricultural Sciences, Ulls väg 16, 750 07 Uppsala, SE

14Agricultural & Biological Engineering Department, University of Florida, Frazier Rogers Hall, Gainesville, FL 32611, USA

15CNR-Ibimet, Via Madonna del Piano 10, 50019 Sesto Fiorentino, IT

16 Department of Agri-food Production and Environmental Sciences - University of Florence, Piazzale delle Cascine 18, 50144 Firenze, IT

17IBOE Soil Ecology, German Research Center for Environmental Health, Ingolstädter Landstraße 1, D 85764 Neuherberg, DE

\*Corresponding author: Gang Zhao, Tel: (+49) 228 732 876, fax: (+49) 228 732 870, email: [gzhao@uni-bonn.de](mailto:gzhao@uni-bonn.de)

# Highlights

Spatial heterogeneity of crop growth environment is ubiquitous over large areas

Often only a limited number of sites are used for simulating regional crop yield

We compare eight sampling schemes for estimating regional mean yield

Precisions of eight schemes are compared across fourteen crop models

The findings can improve the precision of the site-based regional crop modeling

# Abstract

We compared the precision of simple random sampling (SimRS) and seven types of stratified random sampling (StrRS) in estimating regional mean yields for two crops (winter wheat and silage maize) that were simulated by fourteen crop models. We found that the precision gain of StrRS notably varied across stratification methods and crop models. Precision gain for StrRS with even area stratification was positive, stable and consistent across crop models. Stratification with soil had very high gains of precision for twelve crop models, but resulted in negative gains for crop two models. Increasing the sample size up to 200 monotonously decreased the sampling errors for all the sampling schemes. We conclude that even area stratification can modestly but consistently improve the precision in estimating regional mean yields. Using the most influential environmental variable for stratification can notably improve the sampling precision, when the sensitivity behavior of a crop model is known.

**Keywords**: crop model, stratified random sampling, simple random sampling, clustering, up-scaling, model comparison, precision gain

# Introduction

Dynamic crop models are developed for simulation of crop growth and yield in response to various environmental and management conditions at a field scale ([Keating et al., 2003](#_ENREF_36); [van Diepen et al., 1989](#_ENREF_61); [Williams et al., 1989](#_ENREF_65)). To provide summarized information (e.g. mean/total crop production inside a political boundary) for agricultural impact and risk assessment to support policy decisions, crop models need to be applied over large areas. Due to data paucity and computing cost, simulations are typically conducted at a limited number of sample locations across the region, through which results are up-scaled to regional or larger scales ([Ewert et al., 2011](#_ENREF_26)). For example, [Rötter et al. (1995](#_ENREF_51)) chose 18 sites to represent a large watershed, the Rhine basin. [Trnka et al. (2014](#_ENREF_60)) chose 14 sites to represent Europe to simulate the adverse weather events for wheat. [Asseng et al. (2015](#_ENREF_3)) chose 30 sites across the world to simulate temperature effects on global wheat production. The methods used to choose simulation locations, called sampling design, can be used to improve the confidence in the simulation results ([Roleček et al., 2007](#_ENREF_49)).

The geo-referenced samples differ from the ones in classical sampling by the spatial dependence or autocorrelation, e.g. croplands close to each other tend to have similar environmental conditions ([Caeiro et al., 2003](#_ENREF_17)). However, the classical sampling theory is still valid and useful for spatial sampling ([Brus and De Gruijter, 1997](#_ENREF_11); [Brus and DeGruijter, 1993](#_ENREF_13); [De Gruijter and Ter Braak, 1990](#_ENREF_23)). Model-based and design-based are two widely used schemes of sampling ([Cassel et al., 1977](#_ENREF_18); [Wang et al., 2013](#_ENREF_62)). Design-based strategies can be more efficient when probability sampling is possible and sample size is large ([Brus and De Gruijter, 1993](#_ENREF_10)). Two most frequently used design-based strategies are simple random sampling (SimRS) and stratified random sampling (StrRS) ([Hirzel and Guisan, 2002](#_ENREF_33); [Ripley, 2005](#_ENREF_48)). In SimRS, a certain number of independent samples are randomly drawn from the population across a region, each sample with equal probability ([Cochran, 1977](#_ENREF_19)). In StrRS, the entire study area is separated into sub-regions, called strata (or zoning), according to prior information on the samples and then random sampling is applied to each stratum. These two design-based schemes have been widely evaluated in monitoring of natural resources ([Brus, 1994](#_ENREF_9); [De Gruijter et al., 2006](#_ENREF_22)), species distribution modeling ([Stockwell and Peterson, 2002](#_ENREF_56); [Wisz et al., 2008](#_ENREF_68)) and demographic health surveys ([Kumar, 2007](#_ENREF_39), [2009](#_ENREF_40)). In a vegetation survey, [Austin and Heyligers (1989](#_ENREF_5)) found that stratifying the samples by combined information on climate, topographic and lithological characteristics could better represent the environmental variability in the area, especially when the stratification is coupled with well-tuned sampling rules based on aspect and topographic position. [Wang et al. (2002](#_ENREF_64)) found that zoning of the spatial samples based on prior knowledge of the controlling variables could reduce the sample size to achieve the same efficiency in monitoring the area of cultivated land. [Brus (1994](#_ENREF_9)) found that the estimation accuracy can be improved by stratifying the study area based on soil and land use maps when estimating the spatial means of phosphate sorption characteristics. [Brus (2014](#_ENREF_12)) compared three spatial sampling approaches (a design-based, a model-based and a hybrid approach) for regional soil monitoring and found that the performance of the three approaches varied across the different soil properties. [Wang et al. (2010](#_ENREF_63)) found that stratification of study area could reduce the variance of estimators in surveys of non-cultivated land in China.

In these survey and monitoring applications, the prior information that is used to stratify the study areas is normally obtained from other correlated variables or historical survey data. In crop modeling, the results are simulated by the input of environmental variables and management practices, which can be used as prior information to create the strata. Many types of strata or zones including climate zones ([Rötter et al., 2012](#_ENREF_50); [Rötter et al., 1995](#_ENREF_51)), environmental zones ([Metzger et al., 2005](#_ENREF_41); [Olesen et al., 2011](#_ENREF_45)), agro-ecological zones ([Aggarwal, 1993](#_ENREF_1)), and climate-soil zones ([Bryan et al., 2014](#_ENREF_15); [Zhao et al., 2015a](#_ENREF_71)), have been used for regional or global crop modelling studies. However, only very few studies explicitly investigated the precision of these stratification methods and spatial sampling strategies. [Nendel et al. (2013](#_ENREF_44)) showed that one soil profile and weather station were not sufficient to represent the observed mean grain yields of winter wheat in Thuringia, a region in Germany covering more than 16 000 km². By using one soil profile and gridded weather data at 1 km spatial resolution, [Bussel et al. (in review](#_ENREF_16)) evaluated the effects of changing sample size of StrRS on simulations of winter wheat yields under two production conditions, i.e. potential and water-limited in North Rhine-Westphalia. They recommended that detailed soil properties should be included in the simulations to further consolidate the conclusions from their study. To our best knowledge, no study has compared the efficiency of different stratum types (e.g. information used to create the strata) and stratum number for simulating regional crop yields.

This study aims to compare the precisions of SimRS and seven types of StrRS in estimating the regional mean yields for two crops (winter wheat and silage maize). We investigated how the precision, indicated by rooted mean square error (*RMSE*), depends on sample sizes, stratum types, stratum numbers, crop types and crop models. To get the population true mean, we simulated yields for two crops, winter wheat and silage maize, at 1 km spatial resolution with fourteen crop models under water-limited conditions. Seven types and three number (4, 8 and 16) of strata were created by using *k-*means clustering according to coordinates of simulation grid cells (even area), temperature, precipitation, radiation, climate conditions (temperature, precipitation and radiation), soil (soil water holding capacity) and environmental conditions.

# Methods

## Sampling precision

Crop yields of a region (*A*) construct a continuous surface that can be infinitely divided. Due to the computing cost and input data availability, it was not possible to do the simulations for each field of the entire study area. Instead, we divided the region A to 1 km resolution grids and simulated yields for each grid to approximate the surface and the results were treated as the population (*N* = 34168). We sampled the population with a range of sample sizes and sampling schemes to estimate the population mean yield () of the region (*A*) and tested how the sampling precision was influenced by the sample size and scheme choice.

Eight design-based sampling schemes were evaluated, including simple random sampling (SimRS) and seven stratified random sampling (StrRS) with different types of strata created according to corresponding environmental variables. Since the SimRS can be treated as one stratum StrRS, the equations for sampling error estimation for StrRS can be used for SimRS. The population was first divided into subgroups (strata), *N1, N2, …, NL,* which were not overlapping. Samples are independently and randomly drawn from each stratum without replacement. The sample sizes within the strata are denoted by *n1*, *n2*,…, *nL*. The value of *Nh* is known when the strata are created. The suffix *h* denotes the stratum and *i* denotes simulated yields within the stratum. We summarized the symbols used in this study in Table 1.

Table 1 Symbols for simple random sampling and stratified random sampling in this study.

|  |  |  |  |
| --- | --- | --- | --- |
| Symbol | Description | Symbol | Description |
|  | Population size |  | Yield for the *i*th grid/sample in stratum *h* |
|  | True population mean for region *A* |  | Stratum weight of stratum *h* |
|  | Population mean estimator with stratified sampling |  | Sampling fraction in stratum *h* |
|  | Numbers of strata, 4,8 and 16 |  | True mean yield of stratum *h* |
|  | Sub-population sizes in stratum *h* |  | Mean yield estimator in stratum *h* |
|  | Sample size in stratum *h* |  | True variance in stratum *h* |

We denote the population true mean for the whole study region by . The estimated mean with stratified random sampling is which was calculated as

|  |  |
| --- | --- |
|  | (1) |

where .

The estimator is unbiased, since the mean of all possible samples equal to the *true* population mean of stratum *h*. Therefore, is also an unbiased estimator of the population mean of the entire region according to Theorem 5.1 in [Cochran (1977](#_ENREF_19)). To quantify the sampling error of a sampling scheme we use the mean squared error

|  |  |
| --- | --- |
|  | (2) |

Since is unbiased, . The mean squared error equals the variance

|  |  |
| --- | --- |
|  | (3) |

Since the samples in different strata are drawn independently, the covariance between strata equals to 0. Therefore, *MSE* can be calculated as

|  |  |
| --- | --- |
|  | (4) |

According to Theorem 2.2 in [Cochran (1977](#_ENREF_19)), the estimator variance , applied to eq. 4

|  |  |
| --- | --- |
|  | (5) |

To facilitate the interpretation of the results, we took the square root, , to calculate sampling standard error, which is used to measure the sampling precision of a sampling scheme.

At the end, we calculated the precision gain (*PG*) in percentage of different types of stratification by

|  |  |
| --- | --- |
|  | (6) |

To a fair comparison, and always shared the same sampling size.

## Study area

The study area, the state of North Rhine-Westphalia (NRW, 6 E – 9.5 E, 50 N – 52.5 N), is located in the middle-west of Germany (Fig. 1). Flat plain covers 50% of the study area. The topography rises from northwest towards the southeast of the state and merges into Germany's Central Uplands. Agriculture land occupies more than 60 % of the state area. Winter wheat and silage maize are the predominant crops according to the yield report of [Federal Statistical Office (2013](#_ENREF_27)). We simulated crop growth and yields over the entire region (34168 grid cells) without considering the land use in reality. In this way, it represented large heterogeneity of environmental conditions some of which may be suitable for crop growth.

E:\Papers\AgMIP_P2\StudyArea\StudyArea.tif

Fig. 1 The location and terrain of the study area, North Rhine-Westphalia (NRW), Germany (data source: Federal Agency for Cartography and Geodesy, Germany, http://www.bkg.bund.de).

## Climate and soil

Thirty years (1982 – 2011) gridded monthly weather data including maximum, mean, and minimum temperature, sunshine hours, and daily precipitation, at 1 km resolution were obtained from German Meteorological Service ([DWD, 2015](#_ENREF_25)). The data were combined with station-based daily data from more than 200 local weather stations to produce the gridded daily weather data as model inputs. The detailed procedures for fusion of the two data sources were described in [Zhao et al. (2015b](#_ENREF_73)) and [Siebert and Ewert (2012](#_ENREF_52)). A summary of daily mean temperature, annual sum precipitation and annual sum global solar radiation can be found in Fig. 2.

The soil data at a scale of 1: 50000 were obtained from [GDNRW (2001](#_ENREF_30)). The soil data available in vector format were converted to a raster format of 300 m spatial resolution. To convert the soil data from 300 m to 1 km spatial resolution, we took the area-dominant soil types of the 300 m grid cells inside the 1 km x 1 km grid cell and allocated the profile of the dominant soil type to the 1 km resolution grid cells (Fig. 2, e). The profiles of different physical properties sharing the same soil type were considered as unique soil types in the aggregation. The soil water holding capacity (SWHC) and the dominant soil types for each 1 km x 1 km grid cell are shown in Fig. 2 (d and e).

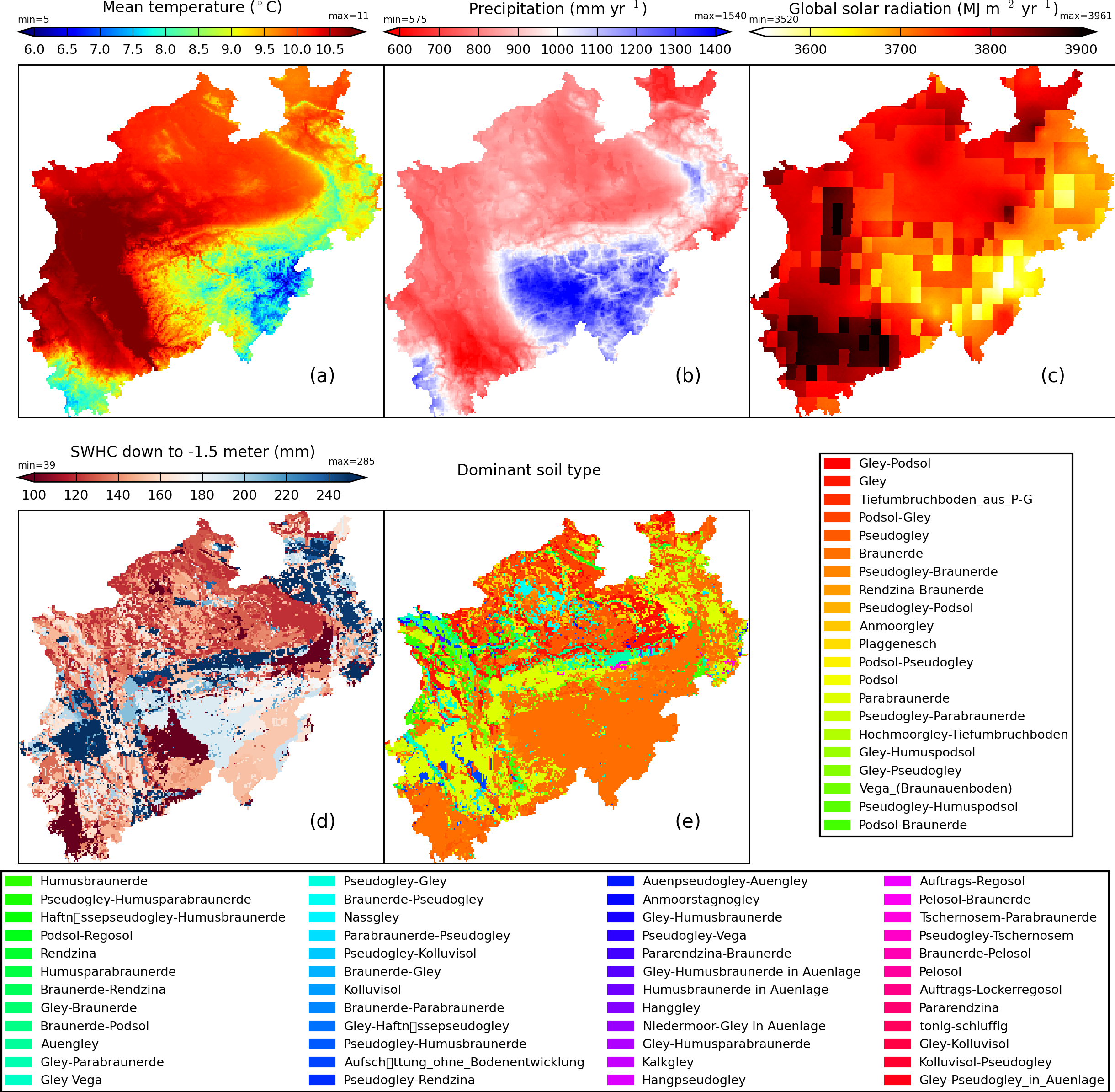


Fig. 2 Summary of the climate and soil data at 1 km spatial resolution for the study area, North Rhine-Westphalia (NRW), Germany. The climate variables include average of daily mean temperature (a), annual sum precipitation (b) and annual sum radiation (c). The soil property variables include soil water holding capacity (SWHC) down to -1.5 m depth (d) and area-dominant soil types at 1 km spatial resolution (e).

## Crop models and simulations set up

To compare sampling schemes across different crop types and crop models, fourteen process-based crop models including AGROC ([Bauer et al., 2012](#_ENREF_6)), APSIM ([Holzworth et al., 2014](#_ENREF_34); [Keating et al., 2003](#_ENREF_36)), APSIM-NWHEAT ([Asseng et al., 2000](#_ENREF_4)), CENTURY ([Kelly et al., 1997](#_ENREF_37); [Parton and Rasmussen, 1994](#_ENREF_47)), CropSyst ([Stöckle et al., 2003](#_ENREF_54); [Stockle et al., 1994](#_ENREF_55)), CoupModel ([Conrad and Fohrer, 2009](#_ENREF_20); [Jansson, 2012](#_ENREF_35)), DailyDayCent ([Del Grosso et al., 2006](#_ENREF_24); [Yeluripati et al., 2009](#_ENREF_69)), EPIC ([Williams et al., 1983](#_ENREF_66); [Williams and Singh, 1995](#_ENREF_67)), Expert-N ([Stenger et al., 1999](#_ENREF_53)), HERMERS ([Kersebaum, 2007](#_ENREF_38)), SIMPLACE<LINTUL> ([Gaiser et al., 2013](#_ENREF_29); [Zhao et al., 2015b](#_ENREF_73)), MCWLA ([Tao et al., 2009](#_ENREF_57); [Tao and Zhang, 2013](#_ENREF_58)), MONICA ([Nendel et al., 2011](#_ENREF_43)) and STICS ([Bergez et al., 2014](#_ENREF_7); [Brisson et al., 2003](#_ENREF_8)), were used to simulated the yields of winter wheat (fourteen models, Fig. 4) and silage maize (ten models, Fig. 5) across NRW (34168 grid cells). The detailed mechanism and configuration of the involved models are provided in Table S1 in the supplementary materials.

Under a water-limited production situation, winter wheat and silage maize were simulated individually in a continuous cropping system. For winter wheat, 400 plants m-2 were sown at the depth of 4 cm on 1st October. For silage maize, 10 plants m-2 were sown at the depth of 6 cm on 20th April. Typical harvest dates for winter wheat (1st Aug) and silage maize (20th Sep) were provided to calibrate the phenology. The historical mean yields (1999 – 2011) of NRW ([Federal Statistical Office, 2013](#_ENREF_27)) for the two crops were used to calibrate the biomass production.

## Stratification and precision quantification

In this study, we aimed to use a limited number of sites to estimate the population mean crop yield, , for NRW. We compared the precisions of eight sampling schemes across models and sample sizes, but not the performances of different models comparing with observations. The population true mean value, , varied across involved crop models.

In the StrRS, prior information on variables including coordinates of grid cells (even area, also named compact geographical stratification in [Brus et al. (1999](#_ENREF_14))), temperature, precipitation, radiation, climate conditions (temperature, precipitation and radiation), soil (SWHC down to -1.5 m) and environmental conditions (climate, soil and terrain), was used to create the strata (Fig. 3). These variable(s) were fed into a *k-means* clustering algorithm and the clustering results were used as strata ([Arthur and Vassilvitskii, 2007](#_ENREF_2)). Since the variables have different orders of magnitude, they were first scaled into zero mean and unit variance. For each stratum type, three numbers of strata (i.e. *L=* 2, 4, and 8) were created by using the corresponding cluster number in the *k-means* clustering. In the SimRS, the entire study area was treated as one stratum (*L* = 1) and the samples were randomly drawn from the population of simulated yields. In the StrRS with seven types of strata, the samples were first evenly allocated to the strata (), from which the samples were randomly drawn. Sample sizes from 2 to 200 were applied to each SimRS. Sample sizes from 2 to 200 only numbers that are divisible by the corresponding stratum number and each stratum had at least two samples applied to the StrRS were evaluated for StrRS. For example, when stratum number was 8, the evaluated sample sizes were 16, 24, …, 200. In StrRS, the weight *wh* equals to the proportion of grid cell number in each stratum to the total grid cell number of the study area. For SimRS, *wh* was set as 1.0, because it was treated as a one stratum StrRS. The true variance, weight and sample size in each stratum were input into eq. 5 to derive the *MSE*, and later *RMSE* for each sampling scheme and sample size. At the end, the precision gains of seven types of StrRS were calculated.

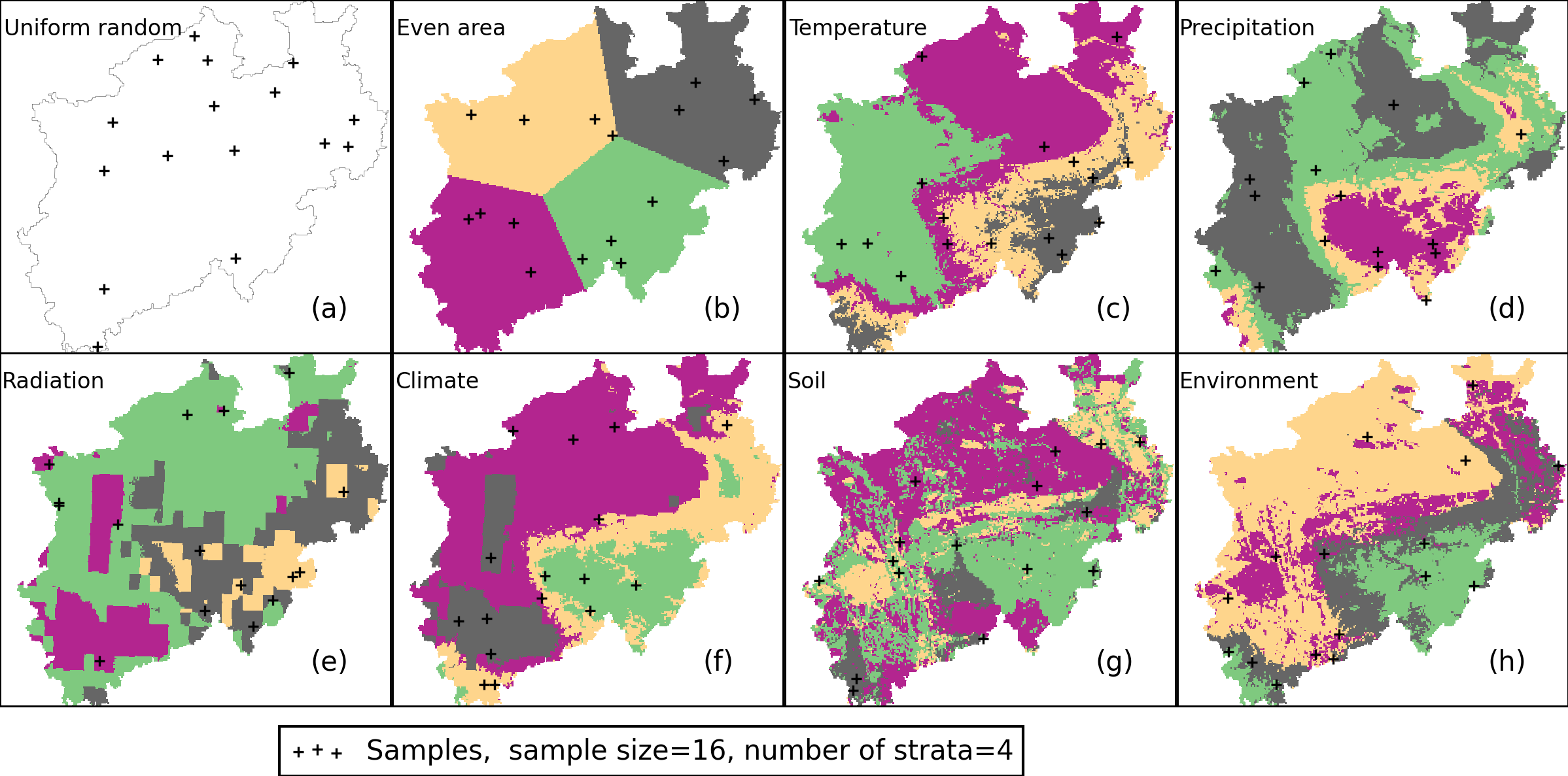


Fig. 3 Illustration of eight sampling schemes: simple random (a), stratification of coordinates, even area (b), stratification of temperature (c), stratification of annual precipitation (d), stratification of annual global radiation (e), stratification of climate conditions (f), stratification of soil (g) and stratification of environmental conditions (climate, soil and terrain) (h). In the stratified random sampling, four different strata are indicated by different colors. The stratifications were created by k-means clustering of the corresponding variables. The stratifications according to climate conditions are clustered according to a combination of three climate variables (temperature, precipitation and radiation). The soil stratifications are clustered according to soil water holding capacity down to -1.5 meter. The stratifications of environmental conditions are clustered according to a combination of climate, soil and elevation.

# Results

## Simulated yields by different models and the models ensemble mean

For winter wheat, the simulated yields (population) by APSIM, APSIM-NWHEAT, EPIC, Expert-N, SIMPLACE<LINTUL> and STICS were higher than by CropSyst and HERMES (Fig. 4, q), which obviously underestimated yields for a larger fraction of the region (Fig. 4, f and i). The majority of the models simulated yields with very high spatial variability. The spatial patterns of simulated crop yields were consistent for the majority of the models. The yields simulated by APSIM, CropSyst, DailyDayCent, EPIC, Expert-N, HERMES, SIMPLACE<LINTUL> and STICS were strongly affected by the soil and showed extremely low yields across the regions with low water holding capacity (SWHC) (Fig. 2 and Fig. 4). The yields simulated by AGROC, APSIM-NWHEAT, CENTURY, CoupModel, MCWLA, and MONICA were less sensitive to SWHC compared to other models (Fig. 4, a, c, d, e, j, and k).

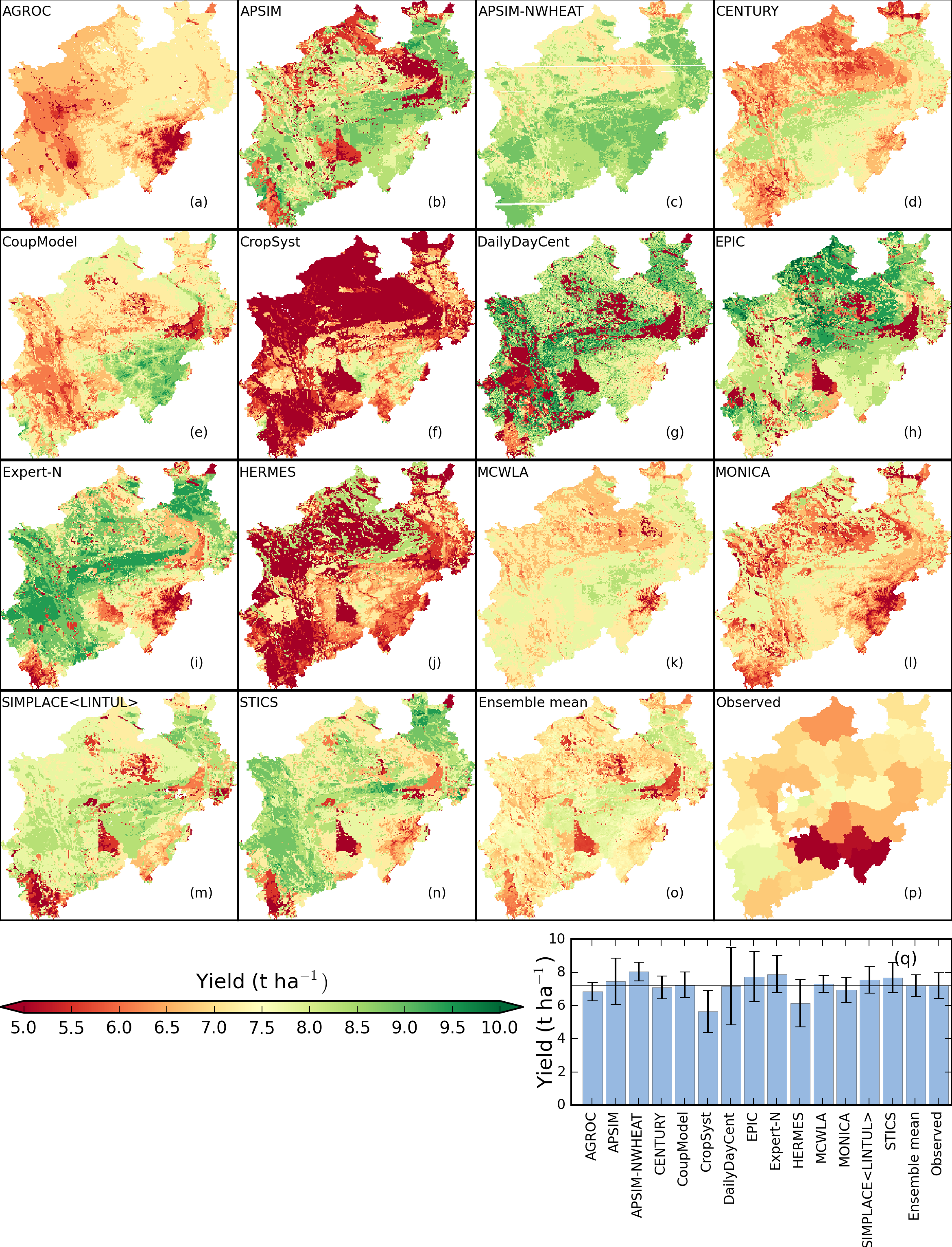


Fig. 4 Spatial distributions of simulated winter wheat yields by fourteen crop models (a – n), the models ensemble mean (o), mean observed yields (1999–2011, 7% of moisture in the reported values) (p), and the mean and standard deviation of wheat yields across grid cells (q). The simulations were conducted under water limited conditions from 1982 to 2011. The observed yields were reported at district levels from 1999 to 2011. In the bar plot (p), the bar height indicates the mean yield across all grid cells and the error bar indicates the standard deviation. The horizontal line indicates the mean observed yield over the study area.

For silage maize, the simulated yields by APSIM, CENTURY, CropSyst, HERMES, MONICA, SIMPLACE<LINTUL> and STICS were higher than by AGROC and EPIC, which showed considerable underestimation (Fig. 5). There were two contrasting spatial patterns across the models. For example, APSIM, CropSyst, MONICA, SIMPLACE<LINTUL> and STICS simulated low yields in the southeastern mountainous regions, while CENTURY, DailyDayCent and HERMES simulated relatively high yields. The yields simulated by the models of APSIM, AGROC, EPIC, HERMES, SIMPLACE<LINTUL> and STICS strongly responded to low SWHC (Fig. 2 and Fig. 5), while yields simulated by CENTURY and MONICA showed a relatively low sensitivity (Fig. 5, c and h).

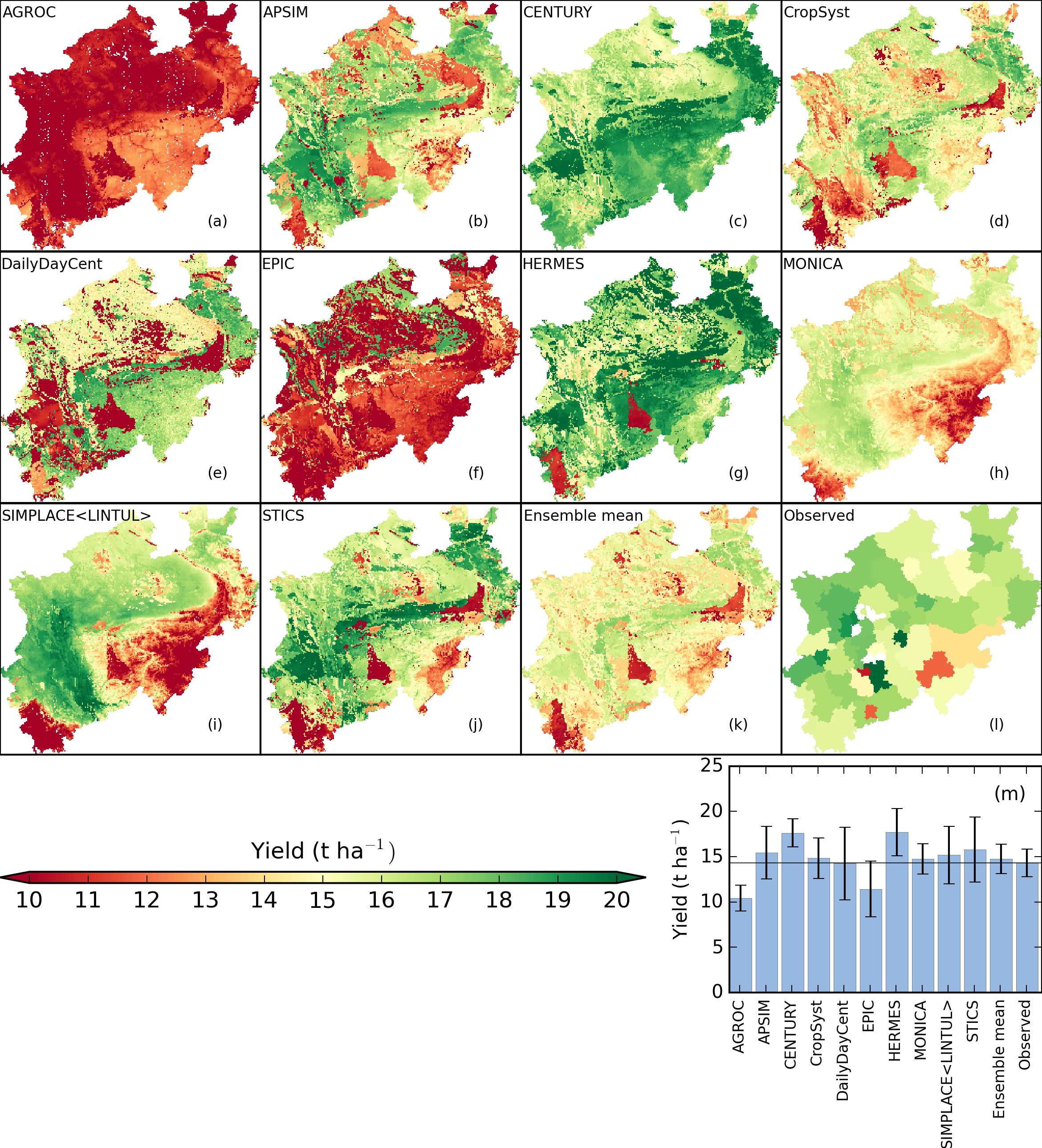


Fig. 5 Spatial distributions of simulated silage maize yields by ten crop models, models ensemble mean, and mean observed yields (1999–2011, 7% of moisture in the reported values) (a-i), and the mean and standard deviation of silage maize yields across grid cells (m). The simulations were conducted under water limited conditions from 1982 to 2011. The observed yields were reported at district levels from 1999 to 2011. In the bar plot (m), the bar height indicates the mean yield across all 34168 grid cells and the error bar indicates the standard deviation. The horizontal line indicates the mean observed yield over the study area.

## Sampling precision across sample sizes

In estimating the regional mean yield for winter wheat with SimRS, sampling standard errors monotonously decreased from 0.750 to 0.075 t ha-1 with the increase of sample size from 2 to 200 (Fig. 6, b). With StrRS (stratum number =8), the sampling standard errors also monotonously decreased across all the seven different stratification methods (Fig. 6, a). The rank of sampling standard errors for eight sampling schemes did not change across sample sizes. StrRS with soil stratification had the smallest errors, and with stratification of climate conditions had the largest ones. The precision with stratification of environmental conditions ranked the second and even area stratification ranked the third.

In estimating the regional mean yield ) for silage maize with SimRS, sampling standard errors monotonously decreased from 1.90 to 0.19 t ha-1 with the increase of sample size from 2 to 200 (Fig. 6, d). With StrRS (stratum number =8), the sampling errors also monotonously decreased across all the seven sampling schemes (Fig. 6, c). Across all the sample sizes, StrRS with soil stratification had the smallest sampling standard errors and highest precision, and stratifications with precipitation, radiation and climate conditions had the largest errors and lest precision. Similar to winter wheat, the rank of sampling standard errors across eight sampling schemes was consistent across sample sizes from 2 to 200. The sampling standard errors of stratification with even area and temperature ranked the third.

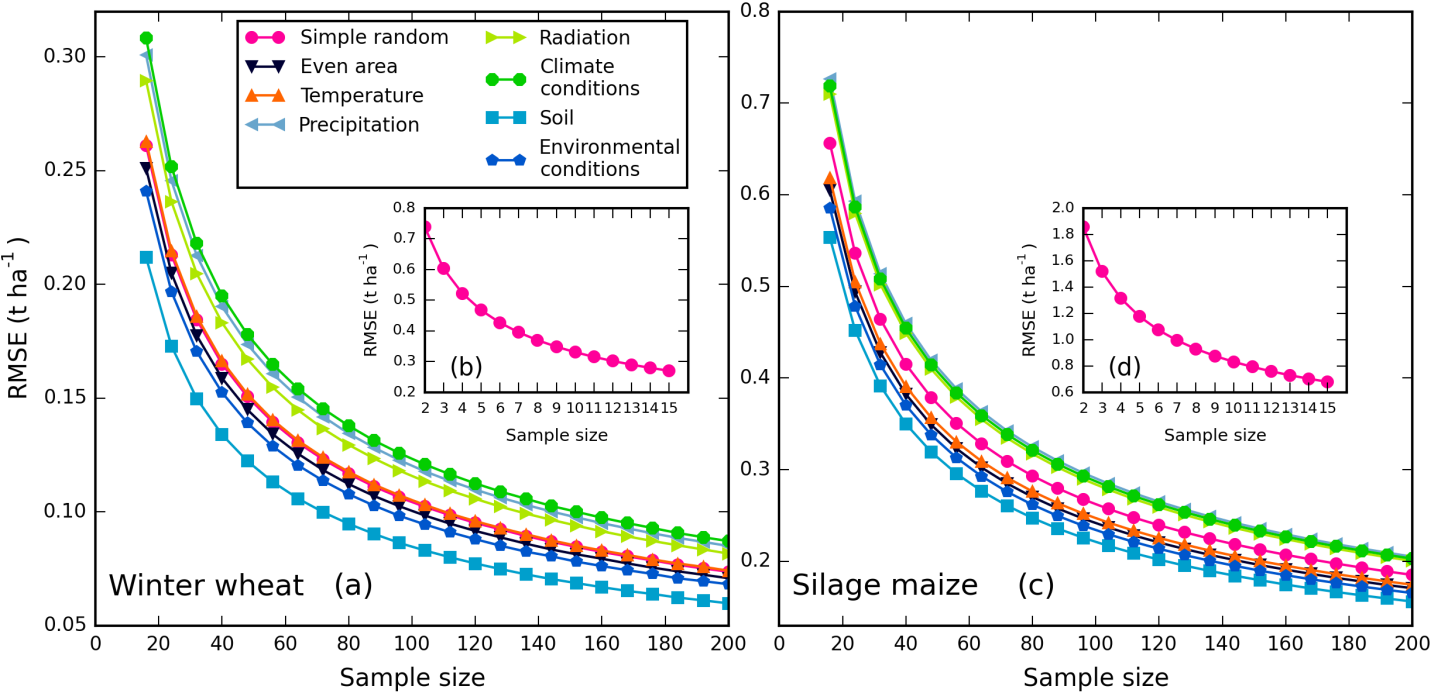


Fig. 6 The variation of sampling standard error (*RMSE*) of eight spatial sampling schemes with the increase of sample sizes (2 to 200 for simple random sampling (SimRS) and 16 to 200 for stratified random sampling (StrRS)) for the mean yield of two crops, winter wheat (a) and silage maize (c). The sub-plots (c) and (d) inside (a) and (b) show data for SimRS from 2 to 15. The sampling standard error of a sampling scheme is indicated by the rooted mean square error (*RMSE*). The stratum number for the StrRS is 8. To use eq. 5 to calculate the sampling standard error, each stratum needs at least two samples so that the smallest sample size for the StrRS is 16.

## The precision gain of stratified random sampling

In estimating the regional mean yields for both crops, StrRS with even area stratification always had positive precision gains (*PG*) and the *PG* increased with enlarging the number of strata from 4 to 16 (Fig. 7, a and b). With 16 strata, the even area stratification can gain precision by 10% (median) for winter wheat and 13% (median) for silage maize. Stratification with temperature, precipitation, radiation and climate conditions did not improve the sampling precision, but rather had negative effects. For both crops, the medians of *PG* for these stratum types were negative, varying from -16% to -2%. Enlarging the number of strata improved the *PG* for stratum type of radiation, but the values were still negative with 16 strata. The stratification of soil and environmental conditions achieved the highest *PG*, especially with large number of strata (8 and 16). The improvement in sampling precision can be as high as 18% (median) for winter wheat and 19% for silage maize. With soil stratification, the *PG* was highest with 8 strata for both crops. With stratification of environmental conditions, the *PG* was highest with 8 strata for winter wheat and 16 for silage maize.

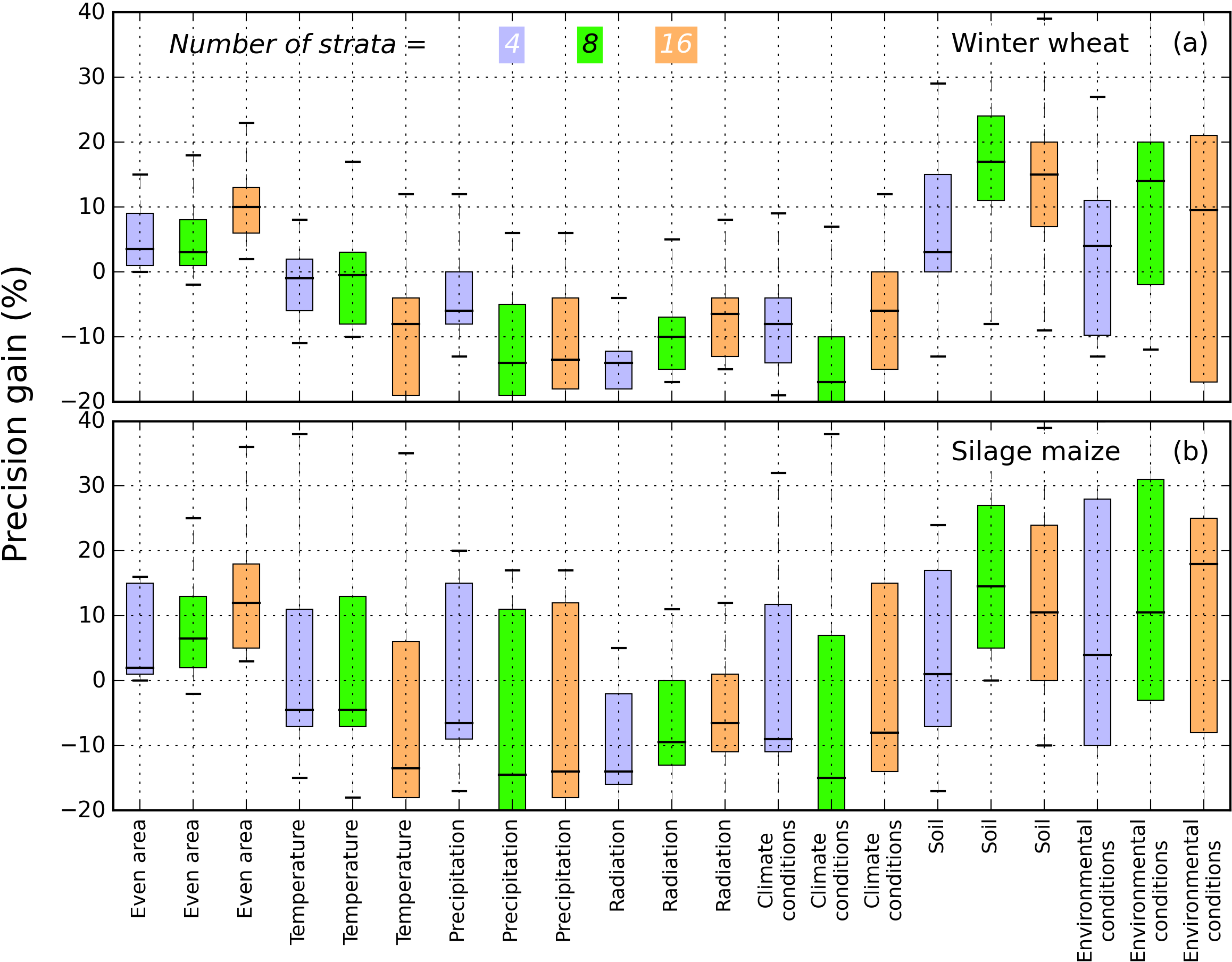


Fig. 7 The precision gain of stratified random sampling comparing for the two crops, winter wheat (a) and silage maize (b). The numbers of strata for the different stratified random sampling schemes are 4, 8 and 16, which are indicated by three different colors as shown in the legend. The variations of precision gain result from crop models (fourteen for winter wheat and ten for silage maize) and sample size (32, 48, …, 192). The edges of the box are the lower hinge (the 25th percentile, Q1) and the upper hinge (the 75th percentile, Q3), and the whiskers extend to Q1 – 1.5 x IQR and Q3 + 1.5 x IQR, IQR = l.5 x (Q3-Q1). The horizontal black lines in the boxes medians of precision gain. The precision gain is calculated with eq. 6.

## Comparing precision gains of the stratified random sampling across crop models

Across all the fourteen crop models, the precision gains (PG) of StrRS with even area stratification were positive (Fig. 8). For even area stratification, the largest PG for winter wheat was with APSIM-NWHEAT (21%) and for silage maize was with SIMPLACE<LINTUL> (36%) (Fig. 8, a). For winter wheat, stratifying the study area with temperature led to negative gains for the majority of the crop models, except for AGROC (48%) and CoupModel (11%). For silage maize, it led to very high gains for SIMPLACE<LINTUL> (54%) and MONICA (35%), but not for other models. Similarly, stratifications with precipitation, radiation and climate conditions also led to negative gains for the majority of models. However, stratification with soil and environmental conditions led to high positive gains for most of the models in estimating the regional mean yields for both crops. Exceptionally, these two stratifications resulted in relatively large negative gains for MONICA (winter wheat) and EPIC (silage maize).

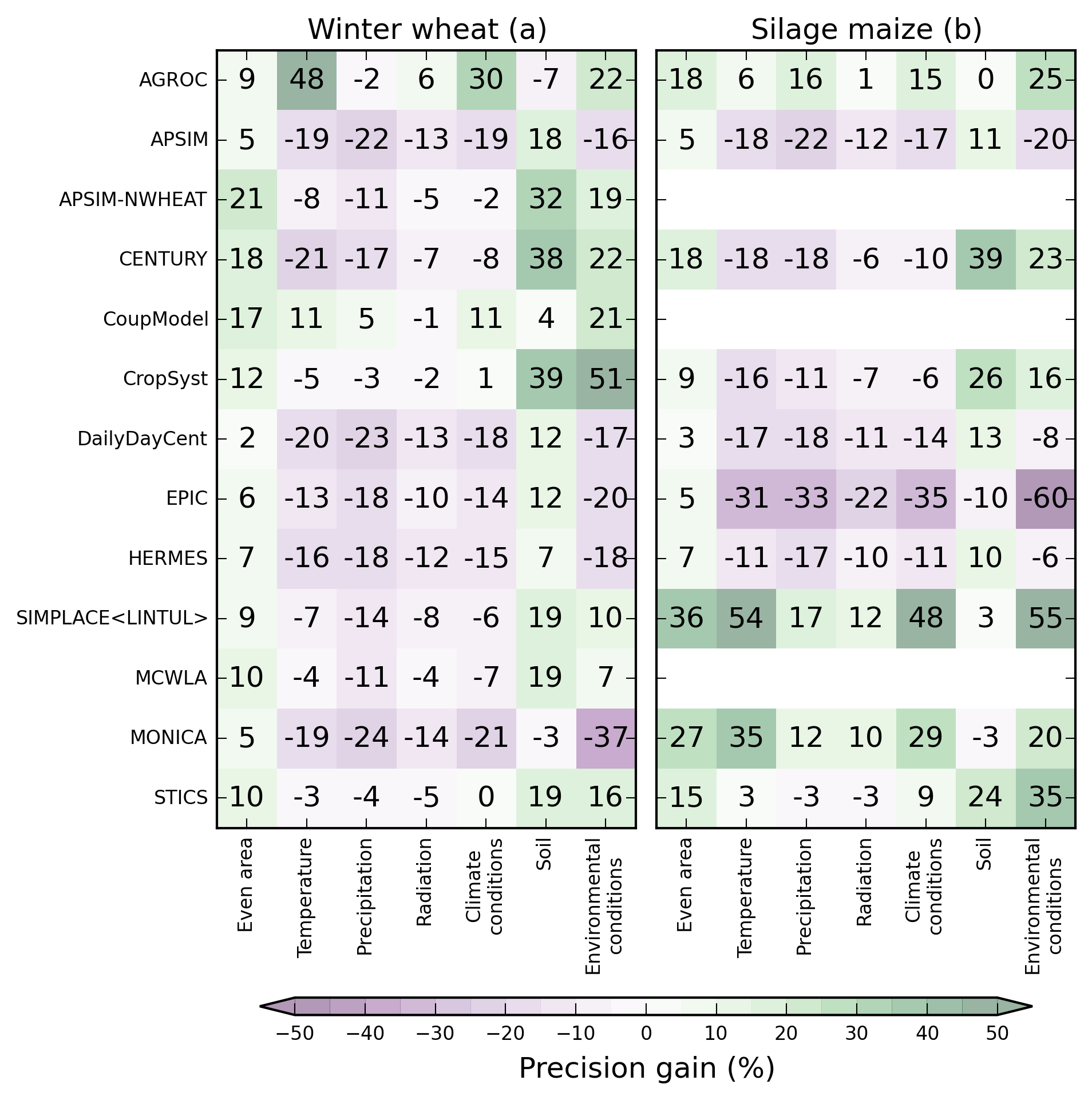


Fig. 8 The precision gain (%) of seven stratified random sampling of fourteen crop models in estimating regional mean yields of the two crops, winter wheat (a, left) and silage maize (b, right). The number of strata for the different stratified random sampling schemes is 16. The sample sizes include 32, 48, …, 192. The precision gain values in the plots are the mean for a crop model in combination with a stratification type across different sample sizes. The precision gain is calculated with eq. 6.

# Discussion

Despite the frequent use of crop models for regional/global change assessments, the access to weather data, soil profiles, crop management information and yield observations is often limited. Since the simulation quality is always affected by the amount and quality of observed data for model input, calibration and validation, elaborating the sampling scheme in order to choose the most representative sites is essential for improving the efficiency and accuracy of crop modeling over large areas ([Wisz et al., 2008](#_ENREF_68)). With an efficient sampling scheme, efforts on experiment implementation and data gathering can target at the most representative sites, thus the cost can be reduced. Although stratification of study area is widely adopted for regional crop modeling, very few previous studies have explicitly tried to quantify the efficiency of and optimize the sampling design. In this study, we tested the precision of eight spatial sampling schemes in estimating the regional mean yields for two crops. The methods used to create the strata in stratified random sampling (StrRS) and the derived knowledge referring to the precisions of different sampling schemes is useful for up-scaling of crop models using a limited number of representative locations. The results from this study can guide the choice of simulation sites when up-scaling crop models by conducting simulations at a limited number of sites.

## Simple random sampling (SimRS) versus stratified random sampling (StrRS)

The assumption behind simple random sampling (SimRS) is that the population is relatively homogeneous and evenly distributed. This is rarely the case for environmental variables across wide geographical areas where spatial heterogeneity and auto-correlation is ubiquitous ([Overmars et al., 2003](#_ENREF_46)). Therefore, the SimRS needs a large sample size to attain the same precision as StrRS (e.g., SimRS reached a precision of 0.22 t ha-1 with a sample size of 32, while StrRS reached with 16, Fig. 6) ([Fortin et al., 1989](#_ENREF_28)). This is particularly true when the environmental gradients are distributed non-randomly in space and similar samples are clumped together. In line with this study, [Mohler (1983](#_ENREF_42)) showed that SimRS could result in truncated response curves for sampling species distribution especially when the extremes of the major environmental gradients were missed in the sampling. Stratifying along these gradients and being particularly careful about sampling the extremes can guarantee an efficient sampling of these outer limits.

To avoid the drawbacks of SimRS, StrRS divides the population to relative homogeneous subgroups so that a smaller sample size is required to the same precision. It fulfills the requirement for that efforts are not wasted for simulating yields across similar environmental conditions for applications of regional crop modeling ([Danz et al., 2005](#_ENREF_21)). According to eq. 5, the sampling standard error (*RMSE*) is determined by the population size, sample size and the true variance in each stratum. Increasing the sample size can continually decrease the sampling standard error but require more resources in gathering the input data and calibration of the models. The computing cost of a large number of simulations can be overcome by taking advantage of advanced hardware and parallel computing ([Zhao et al., 2013](#_ENREF_70)), but long-term records of weather data and soil profile measurements are not available for many locations. Therefore, increasing the sample size is not always suitable to improve the precision of regional crop model applications. Another possible way of increasing the estimation precision is reducing the population variances () in the strata. The smaller the variance in each stratum, the more precision is gained by a StrRS sampling scheme. However, the variances cannot always be reduced by stratification ([Cochran, 1977](#_ENREF_19)). This aspect was proven by the results from this study that many stratification types resulted in negative gains (Fig. 7 and Fig. 8). The reason for the negative gains was that inappropriate prior information was used for the stratification of the population, which did not result in a homogeneous population in the strata. When evenly assigning the samples to strata, many samples were wasted in the relative homogeneous strata and, at the same time, the sample size was too small for many heterogeneous strata. A proportional or optimized allocation of the samples may further improve the precision of the inefficient stratification ([Cochran, 1977](#_ENREF_19)).

For a specific combination of stratum number and sample size, the environmental variables used to create the strata determine the stratum shape, thus the population variance in the strata. If the stratification could reduce the heterogeneity in each stratum, the variance of sampled yields in each stratum can be minimized ([Caeiro et al., 2003](#_ENREF_17)). Therefore, StrRS has a promising potential to improve the regional crop modeling when spatial autocorrelation of the simulated yields is obvious and strong. The results from this study showed that the choice of prior environmental information is critical for the sampling precision of the StrRS. The strata created according to soil data could reduce the sampling error for both of the two crops simulated for most of the models (Fig. 7 and Fig. 8). This is due to the fact that the yields for the two crops were simulated under water-limited conditions and many of the involved crop models are sensitive to the SWHC. The results imply that only the most controlling environmental variable(s) should be used to create the strata in order to minimize the sampling error of StrRS which is also in agreement with the conclusion from [Wang et al. (2002](#_ENREF_64)). Since the sensitivity to different input variables varies across crop models, a type of strata created using prior information of the same variable may have variable outcomes.

We found that using one of the most controlling variables (e.g. SWHC) to create the strata outperformed using many input variables (e.g. climate, soil and terrain) (Fig. 7 and Fig. 6). The reason could be that when the non-influential variable(s) were involved in stratification the simulation results, it cannot reduce the variances of simulation results in the strata. Therefore, choosing one of the most influential environmental variables is recommended to stratify the simulation sites for regional crop modeling.

## Variation of precision gains across crop models

In estimating the mean yield for the two crops, the precision gain of the same StrRS varied across models. The may be due to the different sensitivity behaviors of the simulated results to the environmental variables that were used to create the strata (Fig. 8). If the model does not sensitive to the variable(s) that is used to create the strata, the stratification cannot reduce the heterogeneity in the strata. Therefore, stratification cannot gain any precision, even had negative gains. To verify that the variation of the performance was caused by model sensitivity differences, we conducted a sensitivity analysis of simulated yields to the environmental variables that were used to create the strata (Fig. 9). The results verified our assumption. For example, in estimating the mean yields for maize, SIMPLACE<LINTUL> had a high precision gain with the stratification of temperature (54%), precipitation (17%) and radiation (12%), but only small gain with soil (3%). Correspondingly, the simulated yields of silage maize with SIMPLACE<LINTUL> had a strong correlation with temperature (), precipitation () and radiation (), but only a weak correlation with soil (). Similar correlations between the precision gain a stratum type and the model’s sensitivity to the corresponding variable(s) were also found for other models. Therefore, to be efficient with a StrRS, the strata need to be based on those environmental variables mainly controlling the variability of the output variable of interest.

In the present study, the mean of the simulated yields across all the 1 km grid cells were the population true mean (). Due to the differences in the models’ structure, the environmental variables that should be considered for selection of the simulation sites are crop model specific. Model sensitivity can be analyzed only when the simulations are executed. At the same time, the strata need to be created according to the model sensitivity behaviors. This issue can be solved by a two-step method which has been applied to the natural resource survey ([Guisan et al., 2006](#_ENREF_31)). First, a crop model can be executed across some random locations/sites to study the model sensitivity. Then, the study area can be stratified according to the one of the most sensitive variables derived from the previous steps.

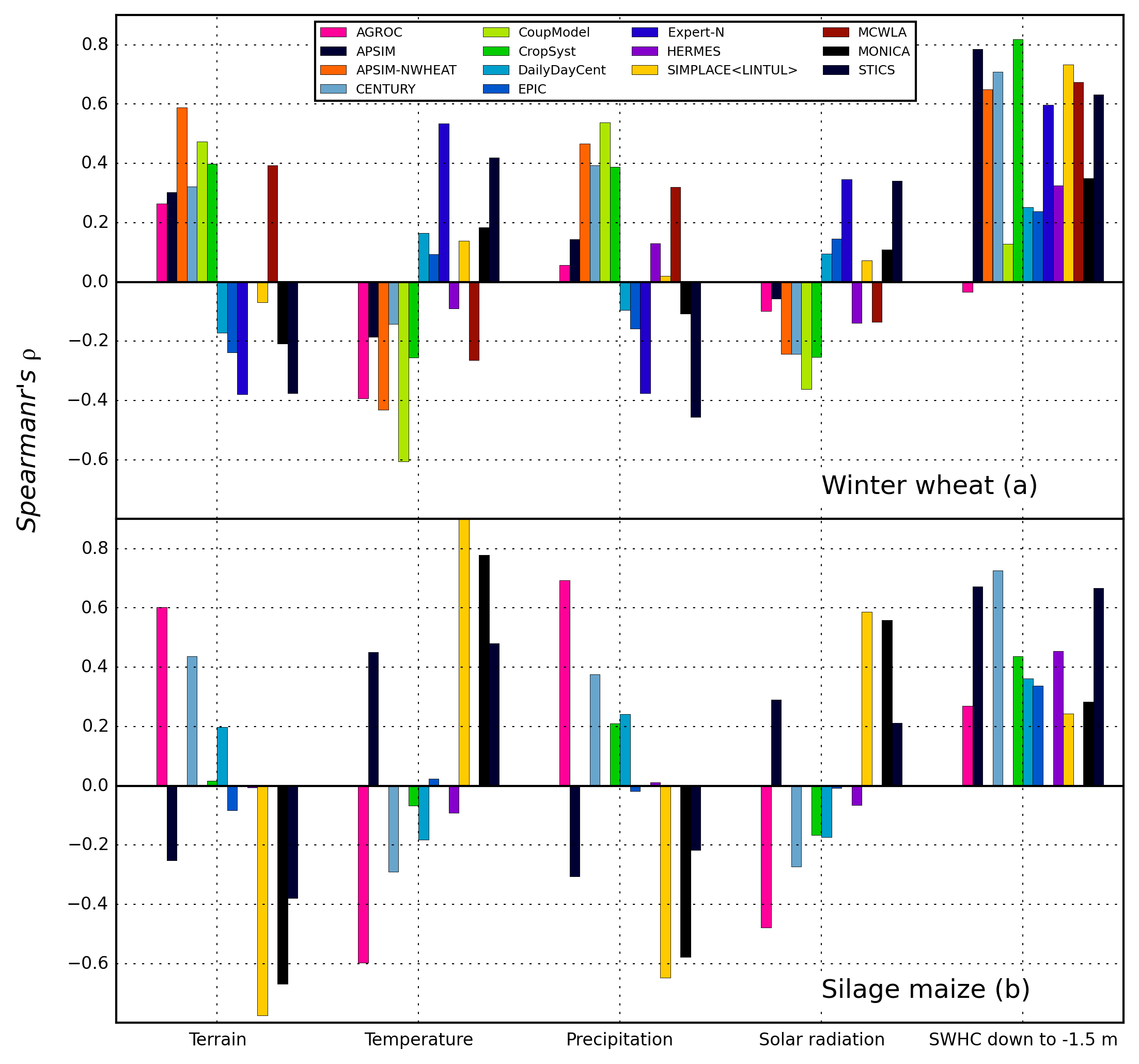


Fig. 9 The correlations between the simulated mean yields by different models and environmental variables, which are used to create the strata for the StrRS. The terrain is the elevation of the grid cell referring to the sea level. Temperature means the daily mean temperature. Precipitation and solar radiation are the mean annual sum of them. SWHC down to -1.5 indicates the soil water holding capacity to -1.5 meter.

## Limitations

Firstly, the crop yields were simulated over the entire region (34168 grid cells) without considering the land use in reality. The large heterogeneity of environmental conditions was represented in the simulations. For large scale studies (e.g. global scale), the crops can be cultivated in severe conditions (e.g. in middle Asia and Africa). Thus, the relatively large heterogeneity of environmental conditions can better test the sampling schemes and inform the sampling strategy for large scale studies. This study focused on the sampling precision for yields simulated under water-limited conditions. In reality, the yields are also influenced by a number of management practices. The interactions between environmental variables and management practices could change the sensitivity of the simulated yields to different environmental variables ([Zhao et al., 2014](#_ENREF_72)). Considering the stress from availability of nitrogen could enhance the yield’s sensitivity to the initial soil organic matter and nitrogen. Secondly, in the StrRS, proportional and optimal allocations of the samples into strata may improve the sampling precision , which was not considered in the present study ([Brus, 1994](#_ENREF_9); [Cochran, 1977](#_ENREF_19)). In the proportional method, the weights for each stratum are equal to each other. It simplifies the way that can be used to evaluate the sampling error, but could miss the hotspots of high or low yields within small areas, especially when the sample size is small. For example, [Hirzel and Guisan (2002](#_ENREF_33)) compared the equal-stratified with the proportional-stratified random sampling scheme in modeling the habitat suitability and found the equal-stratified scheme could achieve a higher efficiency. This conclusion needs to be further validated in modeling crop yield at a regional scale. Furthermore, this study only considered the design-based sampling strategies (e.g. SimRS and StrRS) ([Wang et al., 2002](#_ENREF_64)). The model-based sampling strategies such as the Kriging method have been frequently applied to optimize the spatial sampling ([Hengl et al., 2004](#_ENREF_32); [Theodossiou and Latinopoulos, 2006](#_ENREF_59)). For example, [Brus and De Gruijter (1993](#_ENREF_10)) found that design-based strategies can be more efficient than model-based when probability sampling is possible and sample size is large. A comparison of these two classes of sampling methods can further improve the selection of simulation sites for crop modeling over a large area. Finally, we compared the precision of different sampling schemes in estimating the mean yields over a region, while the spatial structure and inter annual variability of the yields are another aspect that can be of potential interests for some applications ([Fortin et al., 1989](#_ENREF_28)).

# Conclusions

In order to estimate the mean yield of a region, this study investigated how the simulation sites should be distributed along the varying environmental conditions over a large study area. We found that stratified random sampling (StrRS) had a large positive precision gain over the simple random sampling (SimRS) only when the strata were created based on one of the most controlling environmental variables of the respective crop model. Using an even area stratification can modestly improve the sampling precision regardless of crop models and sample sizes. The efficiency of StrRS can be improved by enlarging sample size and stratificaiton of the population or study area based on the most influential environmental variable.

# Acknowledgements

This work was financially supported by the German Federal Ministry of Food and Agriculture (BMEL) through the Federal Office for Agriculture and Food (BLE), grant number 2851ERA01J. The funding support of the Joint Programme Initiative FACCE, MACSUR Knowledge Hub, is greatly acknowledged. S.A. and D.C. were partly funded by a NOAA RISA grant. FT and RR was funded through the FACCE MACSUR project by the Finnish Ministry of Agriculture and Forestry (MMM). EL and EC was supported by The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (contract 220-2007-1218) and by the strategic funding “Soil-Water-Landscape” from the faculty of Natural Resources and Agricultural Sciences (Swedish University of Agricultural Sciences). We are grateful to Professor Per-Erik Jansson (Royal Institute of Technology in Stockholm) for valuable support and recent development of the CoupModel. We are grateful to Gunther Krauss for checking the mathematic equations and terminologies. Comments from three reviewers greatly improved the quality of the manuscript.

# References

Aggarwal, P., 1993. Agro-ecological zoning using crop growth simulation models: characterization of wheat environments of India, Systems approaches for agricultural development. Springer, pp. 97-109.

Arthur, D., Vassilvitskii, S., 2007. k-means++: The advantages of careful seeding, Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics, pp. 1027-1035.

Asseng, S., Ewert, F., Martre, P., Rotter, R.P., Lobell, D.B., Cammarano, D., Kimball, B.A., Ottman, M.J., Wall, G.W., White, J.W., Reynolds, M.P., Alderman, P.D., Prasad, P.V.V., Aggarwal, P.K., Anothai, J., Basso, B., Biernath, C., Challinor, A.J., De Sanctis, G., Doltra, J., Fereres, E., Garcia-Vila, M., Gayler, S., Hoogenboom, G., Hunt, L.A., Izaurralde, R.C., Jabloun, M., Jones, C.D., Kersebaum, K.C., Koehler, A.K., Muller, C., Naresh Kumar, S., Nendel, C., O/'Leary, G., Olesen, J.E., Palosuo, T., Priesack, E., Eyshi Rezaei, E., Ruane, A.C., Semenov, M.A., Shcherbak, I., Stockle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P.J., Waha, K., Wang, E., Wallach, D., Wolf, J., Zhao, Z., Zhu, Y., 2015. Rising temperatures reduce global wheat production. Nature Clim. Change 5(2) 143-147.

Asseng, S., van Keulen, H., Stol, W., 2000. Performance and application of the APSIM Nwheat model in the Netherlands. European Journal of Agronomy 12(1) 37-54.

Austin, M.P., Heyligers, P.C., 1989. Vegetation survey design for conservation: Gradsect sampling of forests in North-eastern New South Wales. Biological Conservation 50(1–4) 13-32.

Bauer, J., Weihermüller, L., Huisman, J., Herbst, M., Graf, A., Sequaris, J., Vereecken, H., 2012. Inverse determination of heterotrophic soil respiration response to temperature and water content under field conditions. Biogeochemistry 108(1-3) 119-134.

Bergez, J.E., Raynal, H., Launay, M., Beaudoin, N., Casellas, E., Caubel, J., Chabrier, P., Coucheney, E., Dury, J., Garcia de Cortazar-Atauri, I., Justes, E., Mary, B., Ripoche, D., Ruget, F., 2014. Evolution of the STICS crop model to tackle new environmental issues: New formalisms and integration in the modelling and simulation platform RECORD. Environmental Modelling & Software 62 370-384.

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., Bussière, F., Cabidoche, Y.M., Cellier, P., Debaeke, P., Gaudillère, J.P., Hénault, C., Maraux, F., Seguin, B., Sinoquet, H., 2003. An overview of the crop model stics. European Journal of Agronomy 18(3–4) 309-332.

Brus, D., 1994. Improving design-based estimation of spatial means by soil map stratification. A case study of phosphate saturation. Geoderma 62(1) 233-246.

Brus, D., De Gruijter, J., 1993. Does kriging really give unbiased and minimum variance predictions of spatial means. J. Soil Sci 44.

Brus, D., De Gruijter, J., 1997. Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). Geoderma 80(1) 1-44.

Brus, D.J., 2014. Statistical sampling approaches for soil monitoring. European Journal of Soil Science 65(6) 779-791.

Brus, D.J., DeGruijter, J.J., 1993. Design‐based versus model‐based estimates of spatial means: Theory and application in environmental soil science. Environmetrics 4(2) 123-152.

Brus, D.J., Spätjens, L.E.E.M., de Gruijter, J.J., 1999. A sampling scheme for estimating the mean extractable phosphorus concentration of fields for environmental regulation. Geoderma 89(1–2) 129-148.

Bryan, B., King, D., Zhao, G., 2014. Influence of management and environment on Australian wheat: information for sustainable intensification and closing yield gaps. Environmental Research Letters 9(4) 044005.

Bussel, L.G.J.v., Ewert, F., Zhao, G., Hoffmann, H., Wallach, D., Constantin, J., Raynal, H., Klein, C., Biernath, C., Heinlein, F., Tao, F., Rötter, R., Cammarano, D., Asseng, S., Elliott, J., Glotter, M., Nendel, C., Kersebaum, K.-C., Specka, X., Basso, B., Baigorria, G.A., Romero, a.C.C., in review. Spatial sampling of weather data for regional crop yield simulations. Agricultural and Forest Meteorology in review.

Caeiro, S., Painho, M., Goovaerts, P., Costa, H., Sousa, S., 2003. Spatial sampling design for sediment quality assessment in estuaries. Environmental Modelling & Software 18(10) 853-859.

Cassel, C.-M., Särndal, C.E., Wretman, J.H., 1977. Foundations of inference in survey sampling. Wiley, New York.

Cochran, W., 1977. G.(1977); Sampling techniques. New York, Wiley and Sons 98 89-100.

Conrad, Y., Fohrer, N., 2009. Modelling of nitrogen leaching under a complex winter wheat and red clover crop rotation in a drained agricultural field. Physics and Chemistry of the Earth, Parts A/B/C 34(8) 530-540.

Danz, N., Regal, R., Niemi, G., Brady, V., Hollenhorst, T., Johnson, L., Host, G., Hanowski, J., Johnston, C., Brown, T., Kingston, J., Kelly, J., 2005. Environmentally stratified sampling design for the development of Great Lakes environmental indicators. Environmental Monitoring and Assessment 102(1-3) 41-65.

De Gruijter, J., Brus, D.J., Bierkens, M.F., Knotters, M., 2006. Sampling for natural resource monitoring. Springer.

De Gruijter, J., Ter Braak, C., 1990. Model-free estimation from spatial samples: a reappraisal of classical sampling theory. Mathematical geology 22(4) 407-415.

Del Grosso, S.J., Parton, W.J., Mosier, A.R., Walsh, M.K., Ojima, D.S., Thornton, P., 2006. DAYCENT national-scale simulations of nitrous oxide emissions from cropped soils in the United States. Journal of environmental quality 35(4) 1451-1460.

DWD, 2015. REGNIE gridded data of daily precipitation. German Meteorological Service (DWD), <http://werdis.dwd.de/werdis/retrieve_data.do?pidpat=de.dwd.hydromet.regnie.daily&toplevel=true>, accessed 08/02/2014.

Ewert, F., van Ittersum, M.K., Heckelei, T., Therond, O., Bezlepkina, I., Andersen, E., 2011. Scale changes and model linking methods for integrated assessment of agri-environmental systems. Agriculture, Ecosystems & Environment 142(1–2) 6-17.

Federal Statistical Office, 2013. Regionaldatenbank Deutschland. https://[www.regionalstatistik.de/genesis/online/logon](http://www.regionalstatistik.de/genesis/online/logon), accessed 08/05/2014.

Fortin, M.-J., Drapeau, P., Legendre, P., 1989. Spatial autocorrelation and sampling design in plant ecology. Vegetatio 83(1-2) 209-222.

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

GDNRW, 2001. AGBK50-Allgemeine Informationen zur Bodenkarte 1 : 50 000 Geologischer Dienst Nordrhein-Westfalen, Arbeitsgruppe BK50. Krefeld, Germany, pp55.

Guisan, A., Broennimann, O., Engler, R., Vust, M., Yoccoz, N.G., Lehmann, A., Zimmermann, N.E., 2006. Using Niche-Based Models to Improve the Sampling of Rare Species

Utilización de Modelos Basados en Nichos para Mejorar el Muestreo de Especies Raras. Conservation Biology 20(2) 501-511.

Hengl, T., Heuvelink, G.B., Stein, A., 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma 120(1) 75-93.

Hirzel, A., Guisan, A., 2002. Which is the optimal sampling strategy for habitat suitability modelling. Ecological Modelling 157(2–3) 331-341.

Holzworth, D.P., Huth, N.I., Zurcher, E.J., Herrmann, N.I., McLean, G., Chenu, K., van Oosterom, E.J., Snow, V., Murphy, C., Moore, A.D., 2014. APSIM–evolution towards a new generation of agricultural systems simulation. Environmental Modelling & Software 62 327-350.

Jansson, P.-E., 2012. CoupModel: model use, calibration, and validation. Transactions of the Asabe 55(4) 1335-1344.

Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Wang, E., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., McCown, R.L., Freebairn, D.M., Smith, C.J., 2003. An overview of APSIM, a model designed for farming systems simulation. European Journal of Agronomy 18(3-4) 267-288.

Kelly, R.H., Parton, W.J., Crocker, G.J., Graced, P.R., Klír, J., Körschens, M., Poulton, P.R., Richter, D.D., 1997. Simulating trends in soil organic carbon in long-term experiments using the century model. Geoderma 81(1–2) 75-90.

Kersebaum, K., 2007. Modelling nitrogen dynamics in soil–crop systems with HERMES, In: Kersebaum, K., Hecker, J.-M., Mirschel, W., Wegehenkel, M. (Eds.), Modelling water and nutrient dynamics in soil–crop systems. Springer Netherlands, pp. 147-160.

Kumar, N., 2007. Spatial sampling design for a demographic and health survey. Population Research and Policy Review 26(5-6) 581-599.

Kumar, N., 2009. An optimal spatial sampling design for intra-urban population exposure assessment. Atmospheric Environment 43(5) 1153-1155.

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

Mohler, C.L., 1983. Effect of sampling pattern on estimation of species distributions along gradients. Vegetatio 54(2) 97-102.

Nendel, C., Berg, M., Kersebaum, K., Mirschel, W., Specka, X., Wegehenkel, M., Wenkel, K., Wieland, R., 2011. The MONICA model: Testing predictability for crop growth, soil moisture and nitrogen dynamics. Ecological Modelling 222(9) 1614-1625.

Nendel, C., Wieland, R., Mirschel, W., Specka, X., Guddat, C., Kersebaum, K.C., 2013. Simulating regional winter wheat yields using input data of different spatial resolution. Field Crops Research 145(0) 67-77.

Olesen, J.E., Trnka, M., Kersebaum, K., Skjelvåg, A., Seguin, B., Peltonen-Sainio, P., Rossi, F., Kozyra, J., Micale, F., 2011. Impacts and adaptation of European crop production systems to climate change. European Journal of Agronomy 34(2) 96-112.

Overmars, K.P., de Koning, G.H.J., Veldkamp, A., 2003. Spatial autocorrelation in multi-scale land use models. Ecological Modelling 164(2–3) 257-270.

Parton, W.J., Rasmussen, P.E., 1994. Long-Term Effects of Crop Management in Wheat-Fallow: II. CENTURY Model Simulations. Soil Sci. Soc. Am. J. 58(2) 530-536.

Ripley, B.D., 2005. Spatial statistics. John Wiley & Sons.

Roleček, J., Chytrý, M., Hájek, M., Lvončík, S., Tichý, L., 2007. Sampling design in large-scale vegetation studies: Do not sacrifice ecological thinking to statistical purism! Folia Geobotanica 42(2) 199-208.

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

Rötter, R.P., Veeneklaas, F.R., van Diepen, C.A., 1995. Impacts of changes in climate and socio-economic factors on land use in the Rhine basin: projections for the decade 2040-49, In: S. Zwerver, R.S.A.R.v.R.M.T.J.K., Berk, M.M. (Eds.), Studies in Environmental Science. Elsevier, pp. 947-950.

Siebert, S., Ewert, F., 2012. Spatio-temporal patterns of phenological development in Germany in relation to temperature and day length. Agricultural and Forest Meteorology 152 44-57.

Stenger, R., Priesack, E., Barkle, G., Sperr, G., 1999. A tool for simulating nitrogen and carbon dynamics in the soil-plant atmosphere system, Proceedings of the technical session, pp. 19-28.

Stöckle, C.O., Donatelli, M., Nelson, R., 2003. CropSyst, a cropping systems simulation model. European Journal of Agronomy 18(3) 289-307.

Stockle, C.O., Martin, S.A., Campbell, G.S., 1994. CropSyst, a cropping systems simulation model: water/nitrogen budgets and crop yield. Agricultural Systems 46(3) 335-359.

Stockwell, D.R.B., Peterson, A.T., 2002. Effects of sample size on accuracy of species distribution models. Ecological Modelling 148(1) 1-13.

Tao, F., Yokozawa, M., 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. Agricultural and Forest Meteorology 149(5) 831-850.

Tao, F., Zhang, Z., 2013. Climate change, wheat productivity and water use in the North China Plain: A new super-ensemble-based probabilistic projection. Agricultural and Forest Meteorology 170 146-165.

Theodossiou, N., Latinopoulos, P., 2006. Evaluation and optimisation of groundwater observation networks using the Kriging methodology. Environmental Modelling & Software 21(7) 991-1000.

Trnka, M., Rötter, R.P., Ruiz-Ramos, M., Kersebaum, K.C., Olesen, J.E., Žalud, Z., Semenov, M.A., 2014. Adverse weather conditions for European wheat production will become more frequent with climate change. Nature Climate Change 4(7) 637-643.

van Diepen, C.A., Wolf, J., van Keulen, H., Rappoldt, C., 1989. WOFOST: a simulation model of crop production. Soil Use and Management 5(1) 16-24.

Wang, J.-F., Jiang, C.-S., Hu, M.-G., Cao, Z.-D., Guo, Y.-S., Li, L.-F., Liu, T.-J., Meng, B., 2013. Design-based spatial sampling: Theory and implementation. Environmental Modelling & Software 40 280-288.

Wang, J., Haining, R., Cao, Z., 2010. Sample surveying to estimate the mean of a heterogeneous surface: reducing the error variance through zoning. International Journal of Geographical Information Science 24(4) 523-543.

Wang, J., Liu, J., Zhuan, D., Li, L., Ge, Y., 2002. Spatial sampling design for monitoring the area of cultivated land. International Journal of Remote Sensing 23(2) 263-284.

Williams, J., Jones, C., Kiniry, J., Spanel, D.A., 1989. The EPIC crop growth model. Transactions of the ASAE 32 497-512.

Williams, J., Renard, K., Dyke, P., 1983. EPIC: A new method for assessing erosion's effect on soil productivity. Journal of soil and water conservation 38(5) 381-383.

Williams, J.R., Singh, V., 1995. The EPIC model. Computer models of watershed hydrology. 909-1000.

Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., Guisan, A., Group, N.P.S.D.W., 2008. Effects of sample size on the performance of species distribution models. Diversity and Distributions 14(5) 763-773.

Yeluripati, J.B., Van Oijen, M., Wattenbach, M., Neftel, A., Ammann, A., Parton, W., Smith, P., 2009. Bayesian calibration as a tool for initialising the carbon pools of dynamic soil models. Soil Biology and Biochemistry 41(12) 2579-2583.

Zhao, G., Bryan, B.A., King, D., Luo, Z., Wang, E., Bende-Michl, U., Song, X., Yu, Q., 2013. Large-scale, high-resolution agricultural systems modeling using a hybrid approach combining grid computing and parallel processing. Environmental Modelling & Software 41(0) 231-238.

Zhao, G., Bryan, B.A., King, D., Luo, Z., Wang, E., Yu, Q., 2015a. Sustainable limits to crop residue harvest for bioenergy: maintaining soil carbon in Australia's agricultural lands. GCB Bioenergy 7(3) 479-487.

Zhao, G., Bryan, B.A., Song, X., 2014. Sensitivity and uncertainty analysis of the APSIM-wheat model: Interactions between cultivar, environmental, and management parameters. Ecological Modelling 279 1-11.

Zhao, G., Siebert, S., Enders, A., Rezaei, E.E., Yan, C., Ewert, F., 2015b. Demand for multi-scale weather data for regional crop modeling. Agricultural and Forest Meteorology 200(0) 156-171.