Key functional soil types explain data aggregation effects on simulated yield, soil carbon, drainage and nitrogen leaching at a regional scale

Elsa Coucheney[a,*], Henrik Eckersten[b], Holger Hoffmann[c], Per-Erik Jansson[d], Thomas Gaiser[c], Franck Ewert[c,e], Elisabet Lewan[a]

[a] Department of Soil and Environment, Swedish University of Agricultural Sciences, Lennart Hjelms väg 9, Uppsala, SE, Sweden
[b] Department of Crop Production Ecology, Swedish University of Agricultural Sciences, Uppsala, Sweden, Germany
[c] Crop Science Group, INRES, University of Bonn, Katzenburgweg 5, 53115 Bonn, Germany
[d] Department of Land and Water Resources Engineering, KTH-Royal Institute of Technology, Stockholm, SE, Sweden
[e] Leibniz Centre for Agricultural Landscape Research (ZALF), Eberswalder Str. 84, 15374 Müncheberg, Germany

ARTICLE INFO

ABSTRACT

The effects of aggregating soil data (DAE) by areal majority of soil mapping units was explored for regional simulations with the soil-vegetation model CoupModel for a region in Germany (North Rhine-Westphalia). DAE were analysed for wheat yield, drainage, soil carbon mineralisation and nitrogen leaching below the root zone. DAE were higher for soil C mineralization and N leaching than for yield and drainage and were strongly related to the presence of specific soils within the study region. These soil types were associated to extreme simulated output variables compared to the mean variable in the region. The spatial aggregation of these key functional soils within sub-regions additionally influenced the DAE. A spatial analysis of their spatial pattern (i.e. their presence/absence, coverage and aggregation) can help in defining the appropriate grid resolution that would minimize the error caused by aggregating soil input data in regional simulations.

1. Introduction

Modelling agricultural production and adaptation to the environment at regional or global scales is receiving much interest in the context of a growing food demand (Tilman et al., 2011) and climate change (Ewert et al., 2015). Two important issues are to identify areas with high yield potential (van Wart et al., 2013) and sustainable management practices that would minimize environmental impacts (e.g. soil degradation, GHGs emissions, nutrient leaching). Process-based soil-crop models describe the flows of mass and energy in the soil-plant-atmosphere system and have been applied and tested in many different contexts, e.g. for CoupModel (Jansson, 2012), STICS (Coucheney et al., 2015), APSIM (Zhang et al., 2012), DNDC (Giltrap et al., 2010), DANUBIA (Lenz-Wiedemann et al., 2010) or CERES, WOFOST, CropSyst, WARM, and SWAP (Confalonieri et al., 2009). As such, they represent valuable tools for predicting agricultural production in diverse agro-environmental contexts (e.g. Jeuffroy et al., 2014) as well as for assessing impacts on the environment; e.g. leaching of nitrates (Conrad and Fohrer, 2009a), changes in soil carbon (Gervois et al., 2008) and GHGs emissions (De Gryze et al., 2011). They are also used to make predictions in response to climate change (e.g. Tubiello et al., 2000) and management changes (e.g. Ng et al., 2000) at the small plot or field scales where input data are considered to be spatially homogeneous. In this context, they are also increasingly applied at regional (e.g. Gaiser et al., 2009) and global scales (Rosenzweig et al., 2014). This raises new challenges related to model input data, calibration and evaluation and the use of different methods of upscaling and downscaling adds new sources of modelling uncertainties (Ewert et al., 2011).

In regional-scale modelling, one major concern is the need to take into account the spatial variability of the environmental conditions (e.g. climate, soils, management practices) used as model inputs. Previous studies showed the effects of input data quantity and quality on model predictions (Grassini et al., 2015) or evaluated model predictions when applied in diverse agro-environmental conditions (Balkovič et al., 2013; Coucheney et al., 2015). Other recent studies have assessed the errors caused by upscaling methods for a range of agro-environmental contexts and models (Hoffmann et al., 2016a; Kuhnert et al., 2016; Van Bussel et al., 2011; Zhao et al., 2015a). These studies are a step further towards the identification and development of scaling methods that minimize these errors for particular climate, soil and management conditions.

Ewert et al. (2011) present different scaling methods from which two main groups can be distinguished: those based on aggregating model outputs (Zhao et al., 2016) and those based on aggregating inputs. In the latter case, the model is usually run for equal-sized grid cells covering the region at a pre-defined resolution (Angulo et al., 2013, 2014; de Wit et al., 2005; Folberth et al., 2012; Hoffmann et al., 2015; Hoffmann et al., 2016a, 2016b; Jégo et al., 2015; Kuhnert et al., 2016; Zhao et al., 2015a). The challenge is then to generate input data for each grid cell (see Zhao et al., 2015b) from data sources available at finer ('upscaling') or coarser ('downscaling') resolutions and which result in the smallest errors in comparison to simulations carried out for each single cell of the grid. Studies of this type have focused mainly on the effects of upscaling climate and soil input data on the predictions for a range of models.

While uncertainties in radiation and precipitation are recognised as a major source of uncertainty for crop yield prediction, de Wit et al. (2005) concluded that average unbiased estimates of weather data are sufficient for predicting yield at the regional scale. Similarly, Angulo et al. (2013, 2014) found that aggregating climate and soil input data from 10 to 100 km had a small effect on barley yields in the south of Finland simulated with four different crop models. Limited data aggregation effects were also obtained when aggregating climate input for simulations of yield and net primary productivity of wheat and maize under water-limited conditions in western Germany using several crop models (Hoffmann et al., 2015; Kuhnert et al., 2016; Zhao et al., 2015a). The DAE related to soil input data were larger in the same region (Hoffmann et al., 2016a). Olesen et al. (2000) and Jégo et al. (2015) found that fine resolution data for both precipitation and soil properties was important for predictions of representative wheat and maize yields in Denmark and southern Quebec respectively. In the latter region, one reason for this was a high pedo-diversity and a strong spatial correlation between soil type and the rainfall data used in the model simulations. DAE are therefore highly dependent on the region. The small effect of weather data aggregation in the German study regions might be due to limited spatial variation in climate combined with a moderate response of modelled yield to those variations (Angulo et al., 2013; Zhao et al., 2015a). Zhao et al. (2015a) further showed that the DAE due to climate aggregation differed among output variables, being positively correlated with the spatial heterogeneity (variation) of the variables concerned. Zhao et al. (2015b) concluded that a high spatial resolution of climate data is desired for regions with high environmental heterogeneity. In line with this, Kuhnert et al. (2016) found a higher impact of climate aggregation on simulated NPP for single years when extreme events such as drought occurred, because model outputs varied spatially more in these years compared with long-term average values. In addition, Hoffmann et al. (2015) showed that the DAE differed among models and identified the climate variables that had the most influence. This study highlighted the importance of also considering the sensitivity of model outputs to the input data (e.g. Hoffmann et al., 2016a). The question arises as to whether the DAE for a specific model and given output variables can be predicted given the spatial distribution of input data and the model sensitivity to the input. For example, variation in altitude was found to be a good proxy of climate DAE for the LINTUL-SIMPLACE model applied in Germany (Zhao et al., 2015b).

Soil properties such as texture, bulk density, porosity and organic matter content strongly influence the soil hydraulic properties (Schaap and Leij, 1998) and biogeochemical processes (Riffaldi et al., 1996). These impact crop growth and its sensitivity to climate under limiting water and nutrient conditions significantly (Kravchenko and Bullock, 2000). Soil variability is therefore one of the most important factors underlying spatial variability in crop yields (Wassenaar et al., 1999). Soil data aggregation may also have a strong impact on simulations of other variables apart from crop yield, such as soil organic carbon stocks (Zhang et al., 2014) or water and N dynamics (Kersebaum and Wenkel, 1998). Furthermore, the importance of soil properties for yield predictions depends not only on the climate (e.g. Timlin et al., 1998) but

also on management practices such as fertilization. For example, Folberth et al. (2016) showed that the variability in yield related to soil type may exceed weather-related variability in scenarios characterized by low fertilization and irrigation amounts. DAE may therefore be larger in regions with a high pedodiversity (i.e. high spatial variation in soil types; Jégo et al., 2015) and in climates or under certain management practices that lead to high water and nitrogen stress.

The present study explores soil DAE on simulated yields of winter wheat, drainage, N leaching and C mineralisation in the region of North Rhine-Westphalia (NRW) in western Germany (Hoffmann et al., 2016a). Simulations were run with a process oriented soil-vegetation model (the CoupModel; Jansson, 2012) for gridded soil data and a spatially uniform climate to ensure that the spatial variability in outputs is related only to variation in soil properties. The soil data was aggregated by selecting the dominant soil at each coarser resolution (i.e. the soil mapping unit covering the areal majority is selected at each coarser scale). The objective of this study was to investigate the contribution of specific soils and their spatial distribution to the DAE for the selected model outputs. We hypothesize that specific combinations of soil properties ('key soils') generate extreme model outputs and that the spatial distribution (e.g. coverage, spatial aggregation) of these critical soils within the region strongly influences the spatial variation of the model outputs and therefore the DAE. To test this hypothesis we propose and apply an approximation of the DAE as a function of these key soils and their spatial coverage in the region. In addition, the influence of the degree of spatial aggregation of these key soils within an area was investigated with respect to four sub-areas of NRW.

2. Material and methods

2.1. Study area

The study area, the state of North Rhine-Westphalia (NRW, 6 E–9.5 E, 50 N–52.5 N, Fig. 1), is located in west-central Germany with a temperate humid climate. Half of the region (34,098 km²) is covered by flat plains and the topography rises from the northeast towards the southeast with a maximal elevation of 843 m. Agricultural land represents > 60% of the area, with winter wheat and silage maize as the main crops.

The whole climate and soil data used in the study can be obtained from Hoffmann et al. (2016b), a brief description is given below.

2.1.1. Climate data

Gridded 1 km × 1 km daily weather data on maximum, minimum and mean temperature, daily rainfall, solar radiation and wind speed over the 29-year period from 1982 to 2011 were obtained by combining daily data from > 200 local weather stations with gridded (1 km) monthly data from the German Meteorological Service (DWD, 2014; see Siebert and Ewert, 2012; Zhao et al., 2015a for precise descriptions). The regional average annual temperature was 9.1 °C and mean annual precipitation was 802 mm with a spatial coefficient of variation of 9.3% and 18% respectively.

2.1.2. Soil data

Gridded 300 m × 300 m soil data (texture, soil layers, depth, C content) were obtained by aggregating mapping units by areal majority using a soil map at a scale of 1:50,000 obtained from the Geological Service North-Rhine Westphalia (Geological Service NRW, 2004), see Angulo et al. (2014) for a more detailed description. The soil data were complemented by soil physical parameters (e.g. water holding capacity) estimated from the texture class by applying pedo-transfer functions developed for German soils (Eckelmann et al., 2005). Topsoil organic carbon and pH were taken from the database FIS StoBo (LANUV, 2014), while the organic carbon content and C:N-ratio of subsoil layers was approximated using pedotransfer functions (Angulo et al., 2014; Eckelmann et al., 2005). Each soil is characterized by a unique set of
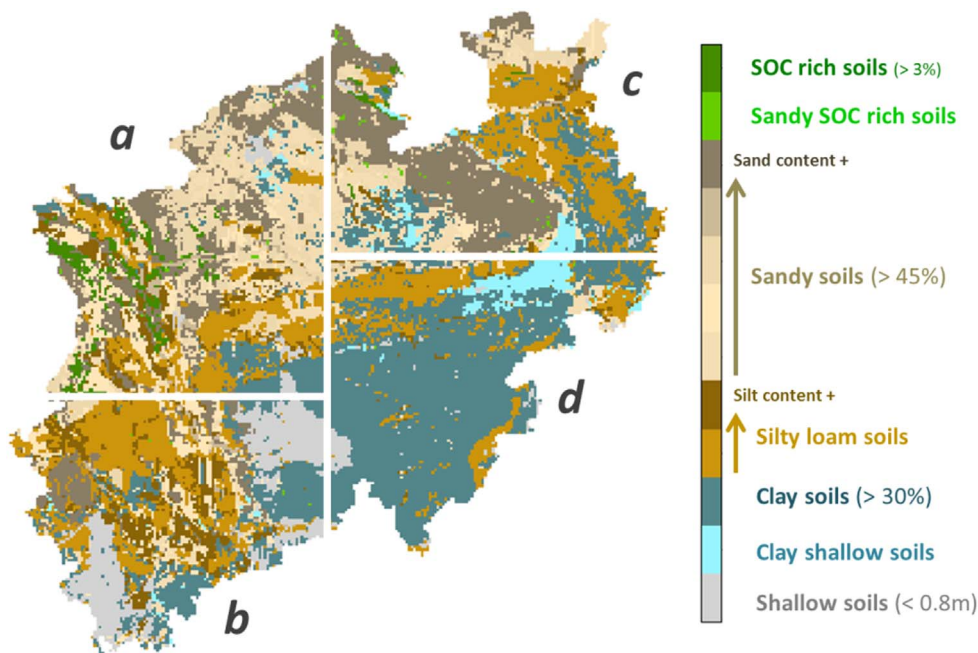
parameters, including its soil class according to the German soil classification system (Eckelmann et al., 2005), soil horizons and their respective thickness, texture, organic carbon content and water holding capacity. The prevailing soils are Cambisols, Luvisols and Stagnosols (FAO key reference soil groups; Hoffmann et al., 2016a). Soil depth was recorded up to 2.3 m and the soil depth varies from 0.4 m to > 2.3 m. The SOC in the topsoil varied between 1.1 and 15.0% with an average value of 2.3%. The soils also showed a wide diversity of texture, from sandy soils (with up to 92% sand in both topsoil and subsoil) to clay soils (with up to 55% clay in topsoil and up to 75% clay in subsoil).

2.2. Model description

CoupModel 5.0 is a dynamic one-dimensional process-based ecosystem model simulating heat, water, carbon (C) and nitrogen (N) transfer in the soil-plant-atmosphere system (Jansson and Moon, 2001; Jansson, 2012; see Jansson and Karlberg, 2013, for a detailed description). The model accounts for soil C and N turnover and N transport linked to the vertical flow of water and heat, in a layered soil profile. Soil water and heat flows are calculated with the Darcy and Fourier equations. Soil C and N are partitioned between two compartments classified as litter (undecomposed crop residues, dead roots and microbial biomass) and humus. The model simulates soil temperature, soil moisture and water flow, and plant development and growth in response to daily weather (e.g. solar radiation, rain, wind speed and mean air temperature). The simulated physical conditions in the soil in turn regulate evapotranspiration, decomposition of soil organic matter and nitrogen dynamics. Feedbacks between the biotic system and the physical environment are simulated on a daily basis.

The main model options and parameters used in this study are presented in Appendix 1 and are the same as those used for the CoupModel simulations in the multi-model scaling exercise presented by Hoffmann et al. (2016a). Most parameters were assumed to be spatially uniform, either set to their default values (Jansson and Karlberg, 2013), taken from previous applications representing arable land in North Europe (Conrad and Fohrer, 2009a; Gustafsson et al., 2004) or in a few cases adjusted to fit mean regional observations (see Appendix 1 on model parameterisation). Other parameters varied spatially depending on soil properties and based on the information in the NRW soil database.

2.3. Model application

2.3.1. Reference simulations with high grid resolution inputs

Soil data was spatially aggregated in order to obtain grids of 1 km resolution by selecting the dominant soil present at 300 m resolution (Hoffmann et al., 2016a). Thereafter, reference model simulations were run with the resultant soil input data at 1 km resolution. For each of these 34,168 grid cells, one simulation was performed representing a monoculture of winter wheat during a 29-year period (1982–2010). The regional climate data series was used as driving data for all model runs. The same management was applied in all cells, with the winter wheat sown every year on October 1st, and fertilised three times with mineral fertilisers according to amounts typically applied in the region (Table in Appendix 1). The harvest date was simulated by the model. Straw was removed but stubble was left on the field (10% of aboveground biomass plus roots). Ploughing occurred on August 31st.

2.3.2. Aggregation of soil data for coarse grids simulations

Soil data was further spatially aggregated in order to obtain grids of 10, 25, 50 and 100 km resolution which correspond to 410, 80, 25 and 9 grid cells, respectively. The model was run for each of the coarser grid cells in a similar way as explained above. This means that the same soils were present at the different resolutions but in different proportions and that many of the soils present at a resolution of 1 km disappeared at coarser resolutions. Fig. 2 shows the areal coverage of the dominant soils at different resolutions (note that only the 223 most abundant dominant soils at a resolution of 1 km are shown, whereas a total of 2646 dominant soils were found at this resolution). The number of dominant soils at resolutions 10, 25, 50 and 100 km were 223, 65, 23 and 8 respectively. Fig. 2 further illustrates that the most common dominant soil at one resolution does not necessarily remain the most common after aggregation to a coarser resolution.

2.3.3. Selected model output variables

Four typical output variables representing the soil-plant system were used to evaluate the DAE. These were (i) the annual drainage (mm) and (ii) annual nitrate leaching (kg N ha^{-1}) at a depth of 1.5 m, or in the case of shallower soils, at the bottom of the soil profile, (iii) annual soil organic C mineralisation in the whole soil profile (g C m^{-2}) and (iv) winter wheat grain yield (estimated from the grain C content at
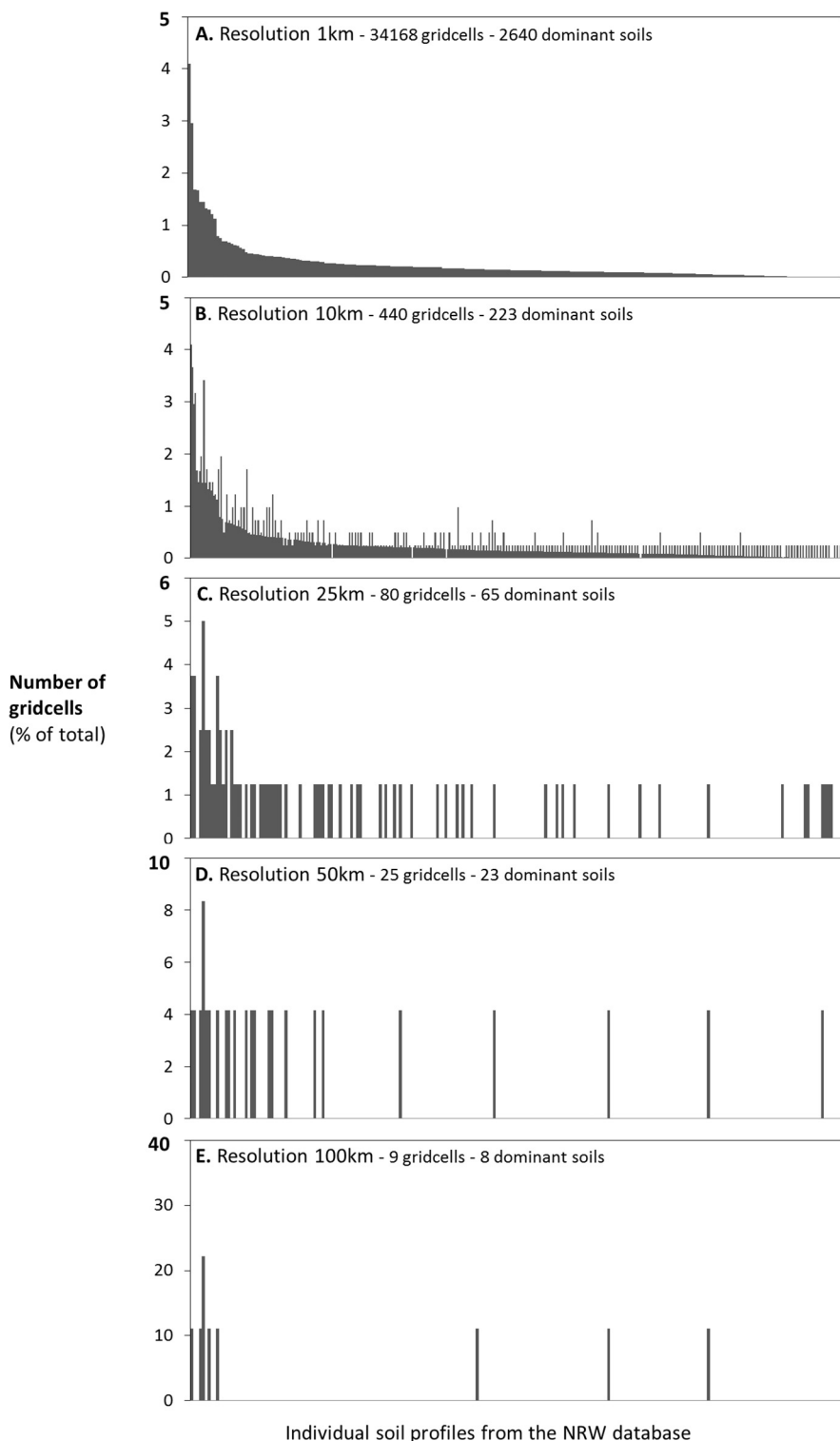
Fig. 2. Number of grid cells covered by individual soil profiles present in the NRW soil database at different resolutions when aggregating by areal majority. Each soil profile represents a soil mapping unit. Soils are ordered from highest spatial coverage (left) to lowest spatial coverage (right) at 1 km resolution. Soils only present at resolution 1 km are not represented.

the date of harvest; t DW ha^{-1}). For each of these outputs, we calculated the regional mean and the between-grid cell variation (CV, %) of the period average (i.e. mean of 29 years).

2.4. Identification of functional soil types

To evaluate if the DAE were related to the presence and spatial distribution of sensitive *functional soil types* (Cécillon and Barré, 2015), we first allocated soils of the NRW database to clusters of soils sharing similar combinations of soil properties (Section 2.4.1). Second, a model

sensitivity analysis for the different soils present in the region was used to identify which of these clusters were associated with extreme values of simulated yield, drainage, C mineralisation or N leaching. Hereafter, these key functional soil types are denoted as "*key soils*" (Section 2.4.2).

2.4.1. Allocating soils to clusters of similar soils

Soils sharing similar combinations of soil properties were allocated to clusters by a *k-means* clustering analysis. To do so, four soil properties in three distinct layers of the soil profile were considered: thickness (m), texture (clay and sand content, %) and organic C content

(%). The three soil layers were topsoil (the uppermost layer recorded in the database), the root zone (zone between the base of the topsoil and the maximal rooting depth of the plant, here 1.5 m) and subsoil (below the root zone). The *k-means* clustering method was based on the dissimilarities (calculated as 'Euclidean distances') of the 12 soil variables (i.e. 4 soil properties × 3 soil layers) standardized to a mean value of zero and a standard deviation of 1. This was done using the '*k-means*' function of the package '*stats*' of R (download at https://www.r-project.org/) and the Hartigan-Wong algorithm (Hartigan and Wong, 1979). To pre-select the number of clusters, we used the subjective elbow method (Thorndike, 1953), which evaluates the improvement of the within-cluster sum of squares (WSS) when one cluster is added. In other words, we did 99 preliminary *k-means* clustering with an increasing number of clusters, from 2 to 100, and the resulting WSS values were plotted against the number of clusters (results not shown). The WSS decreased much more slowly when the number of clusters increased above 10. The number of clusters was then fixed to 10 and the algorithm was run with a maximum iteration of 20 and using the best possible result after 500 random starts.

2.4.2. Identification of key functional soil types

To identify soil types (*key soils*) to which the simulated outputs are sensitive, the model was run once with each of the soils present in the NRW region (i.e. *n* = 2646 simulations). The means and standard deviations of the four simulated output variables ($\overline{S_{allsoils}}$ and $\sigma_{S_{allsoils}}$) were compared with the corresponding values calculated for each soil cluster ($\overline{S_{scx}}$ and σ_{S_x}). This comparison allowed (1) estimation of the deviation from the average of 'all soils' simulations ($D_{scx} = \overline{S_{scx}} - \overline{S_{allsoils}}$) and (2) identification of highly variable clusters with the condition $\sigma_{S_x} > \sigma_{S_{allsoils}}$. Clusters for which D_{scx} was more than twice the standard deviation of the 'all soils' simulations were considered to be associated with extreme simulated values and to belong to a group of '*key soils*' (labelled '*ks*') for the variable concerned; D_{scx} for these '*key soils*' is hereafter denoted D_{ks}. This was done separately for each selected simulated variable. There can be more than one group of *key soils* per variable (or none) and different variables can share the same *key soils*. The influence of the *key soils* can be either negative or positive.

Highly variable clusters (i.e. $\sigma_{S_x} > \sigma_{S_{allsoils}}$) were further divided into two sub-clusters based on the soil input variables (if any) that were significantly correlated to the highly variable simulated output. The resulting new clusters were then tested, in the way described above, to determine whether they belonged to a group of '*key soils*' or not. All soil clusters that did not deviate from the '*all soils average*' were considered as redundant soils and are hereafter denoted '*non-key soils*'.

2.5. Soil spatial diversity and spatial pattern of key soils

To test the effects of soil spatial patterns and diversity on DAE, the NRW region was divided into four approximately equally-sized sub-areas ('*a*', '*b*', '*c*', '*d*'; Fig. 1), by north-south and east-west transects through the approximate centre of the NRW.

We calculated three different normalized Shannon indices (Ramezani, 2012) of soil diversity at 1 km resolution in the NRW region and in the four sub-areas, based on the relative distributions of (1) *all soils* defined in the database (*n* = 2646; H_s), (2) of the soil clusters (H_c) and (3) of the *key soils* versus *non-key soils* (H_{ks} which is referred to hereafter as the uniformity index for *key* and *non-key soils*) by:

$$H = -\sum_{i=1}^{S} p_i\, ln(p_i)/ln(S) \tag{1}$$

where *S* is the number of soils or soil clusters (*S* = 2 when considering *key soils* versus *non-key soils*) and p_i the areal coverage of the soil *i* (or cluster or *key/non-key soils*) in the NRW region or in the sub-areas. This normalized index, which varies between 0 and 1, depends uniquely on the relative distribution of soils and not on their total coverage. It is

close to zero if one soil dominates the whole area, whereas it is close to 1 if all soils are equally represented. We used this index to compare soil patterns in the four sub-areas with each other and with that of the whole region. We also investigated the extent to which these indices could be related to the DAE of model simulations.

The spatial distribution of *key soils* in NRW and the four sub-areas at the reference resolution was characterized by (i) the total spatial coverage (%) estimated as the fraction of 1 km² grid cells covered by *key soils*, and (ii) by spatial aggregation assessed by Joint Count Analysis (JCA, Cliff and Ord, 1973) using the function '*joincount.test*' of package '*spdep*' of R (Bivan et al., https://cran.r-project.org/web/packages/spdep/index.html). This function quantifies the extent to which the spatial pattern of soils is aggregated. If it is not significant, then the *key soils* are considered to be randomly distributed within the area. The grid cells covering the region were colour-coded grey (G) for the *key soils* and white (W) for the *non-key soils* (Fig. 3). The number of G-G (i.e. one edge of a grey-coloured 1 km² grid has a neighbour 1 km² grid that also is grey), W-W and G-W connections were determined and compared with the number of connections that would be expected with a random distribution of G and W. When the pattern was significantly aggregated (*P* < 0.05), we further quantified the degree of spatial aggregation with the normalized aggregation index (*AI*) defined by He et al. (2000). It varies from zero (no spatial aggregation) to one (maximal degree of spatial aggregation).

The spatial coverage of *key soils* and *non-key soils* (respectively Grey and White in Fig. 3) at coarser resolutions than the reference resolution was also assessed. For an area covered by *key soils* at a coarse resolution (Fig. 3A), we distinguished the part of this area that was already covered by *key soils* at a resolution of 1 km (G to G) from the remaining part that was covered by *non-key soils* (W to G, a_0 in Fig. 3). Similarly, for an area covered by *non-key soils* at the coarser resolution (Fig. 3B), we distinguished the part of this area that was converted to *non-key soils* from *key soils* (G to W; a_{ks} in Fig. 3).

2.6. Evaluation of DAE

The DAE was quantified in terms of the discrepancies between model outputs from simulations with soil data aggregated to 10, 25, 50 or 100 km ('*coarse resolutions*'; S_c) and the outputs from the reference simulations with soil data at 1 km resolution ('*true values*'; S_r; Zhao et al., 2015a). This was done with two different metrics.

Firstly, a regional bias for the simulated variables at coarser resolutions was calculated as the relative difference between their arithmetic mean values ($\overline{S_c}$) and the corresponding mean values for the reference simulations at 1 km resolution ($\overline{S_r}$):

$$Bias = ((\overline{S_c} - \overline{S_r})/\overline{S_r})\,100 \tag{2}$$

Secondly, a regional mean error of the simulations at coarse resolutions compared with the reference simulations was calculated as the normalized and area-weighted Root Mean Square Error (*rRMSE*; %, see Eqs. (3)–(7) in Table 1a of Appendix 2). In contrast to the regional bias, this second DAE metric accounts for spatial differences in aggregation errors among coarse grid cells and is also not affected by compensating effects between cells.

2.7. Approximation of the DAE based on key soils

An approximation of the DAE expressed as the regional mean error was made based on the spatial coverage of *key soils* and *non-key soils* at the coarse resolution and at the reference resolution (*rRMSE*** see Eqs. (3) & (8)–(11) in Table 1b of Appendix 2 and Fig. 3). For this, we distinguished two types of errors for a single coarse grid cell, denoted *e1* and *e2* (Eqs. (10a) & (10b)), depending on whether the dominant soil of the coarse grid was a *key soil* (*ks*; grey in Fig. 3A) or a *non-key soil* (*s0*; white in Fig. 3B). To calculate *rRMSE***, any grid cell covered by a *non-key soil* is allocated the mean simulated value for *all soils* ($\overline{S_{allsoils}}$), while
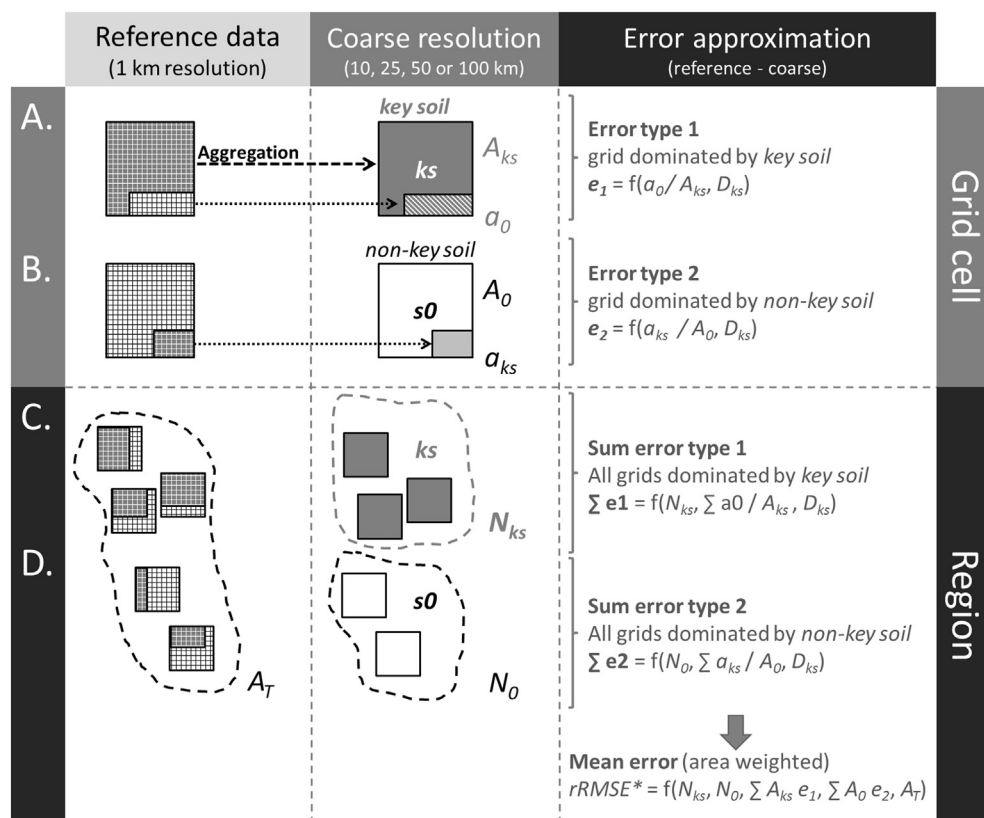
Fig. 3. Approximation of the DAE defined as the Regional Mean Error and calculated with *rRMSE* (Sections 2.6 & 2.7 and Table 1a & 1b in Appendix 2) based on the spatial coverage of *key* soils ('*ks*'; grey grids) and *non-key soils* ('*s0*'; white grids) with respect to one coarse grid cell (A. and B.) or to the whole region (C. and D.).

- A_c is the area of one coarse grid and is denoted A_{ks} when *key soil* is dominant (A.) and A_0 when *non-key soil* is dominant (B.).
- D_{ks} is the difference in a simulated variable between the value associated with a *key soil* and the average value of *all soils* ($S_{allsoils}$).
- a_0 and a_{ks} are the area covered by *non-key soils* inside a coarse grid '*ks*' (A.) and by *key soils* inside a coarse grid '*s0*' (B.) respectively.
- A_T is the total area of the NRW region and N_{ks} and N_0 are the numbers of coarse grids in the region where *key soils* are dominant (C.) and where *non-key soils* are dominant (D.), respectively.

any grid cell covered by a *key soil* is allocated the value ($\overline{S_{allsoils}} + D_{ks}$) (Eqs. (10)–(11)), where D_{ks} (see Section 2.4.2) represents the absolute aggregation error associated with areas that are either converted from *non-key soils* to *key soils* (a_0; Fig. 3A) or from *key* to *non-key soils* (a_{ks}; Fig. 3B). Similarly, the average value of an output variable at the 1 km resolution within one coarse grid cell ($\overline{S_{r,c}}$) was approximated as a function of the fractions of 1 km^2 grid cells that were *non-key* and *key* soils (a_0 and a_{ks} respectively; Eqs. (10)–(11) in the Appendix 2). Finally, for simulated variables with more than one group of *key soils*, the errors caused by the different *key soils* were summed in the calculation of the *RMSE**.

3. Results and discussion

3.1. Reference simulations with inputs at 1 km^2 resolution

The CoupModel generated outputs that were representative of the multi-model ensemble previously applied within the NRW region (Grosz et al., 2017; Haas et al., 2015; Hoffmann et al., 2015, 2016a; Kuhnert et al., 2016; Zhao et al., 2015a). The regional average yield of winter wheat for the simulated period was 7.1 DM/ha (Table 1) and the median value was 7.3 t DM/ha, which is within the range of observed values in the region and represents a simulated yield gap of 20%. The coefficient of variation in simulated yields for 1 km^2 grid cells caused

Table 1
Regional mean of the period average and the within grid cells variation, expressed as a CV, % of annual simulated variables over the 29-years period (1982–2012) obtained at resolution 1 km, as well as data aggregation effects (DAE) for the respective variables (crop yield, water drainage, C mineralization and N leaching) at resolutions 10, 25, 50 and 100 km in NRW. Absolute regional bias > 3.5 % are highlighted in bold.

Variables	29-years period average		Regional Bias (Bias) Eq. 2 in 2.6				Mean Regional Error (rRMSE) Eqs. 3-7 in Appendix 2				
	Regional Mean	Spatial var.	10 km	25 km	50 km	100 km	10 km	25 km	50 km	100 km	Mean
Yield	7.1 t ha⁻¹	8 %	0.2 %	-0.8 %	1.1 %	-1.0 %	4.4 %	5.9 %	7.0 %	8.5 %	6.5 %
Drainage	429 mm	14 %	0.4 %	0.2 %	0.4 %	1.5 %	7.3 %	5.9 %	7.4 %	10.0 %	7.7 %
C mineralization	125 g m⁻²	21 %	-1.5 %	**-4.3 %**	**3.8 %**	-0.6 %	12.0 %	14.2 %	21.7 %	11.1 %	14.8 %
N leaching	83 kg ha⁻¹	29 %	-1.9 %	-1.0 %	**3.6 %**	**5.6 %**	19.0 %	22.8 %	33.3 %	20.2 %	23.8 %
					Mean (rRMSE)		10.7 %	12.2	17.3	12.5	13.2 %

by soil variability was 8%, which is smallest of the four simulated output variables (Table 1).

The regional annual average drainage for the period was 429 mm which represents about half the average annual precipitation in the region. The variation among 1 km² cells (CV of 14%, Table 1) was slightly higher than for the yield. The variations in C mineralisation and N leaching (CV > 20%, Table 1) were much higher than for yield and drainage due to their strong dependence on soil properties, especially the SOC content. The regional annual C mineralisation was in the range of observed values for conventional croplands on Cambisols in a temperate climate (Merino et al., 2004) and represented an annual loss of about 0.8 ± 0.2% of the initial soil C storage (125 g m⁻², Table 1). Mean annual N leaching (83 kg ha⁻¹, Table 1) represented 26 ± 8% of the total amount of N added by fertilisers in the model (317 kg ha⁻¹, Table Appendix 1).

3.2. Functional soil types

3.2.1. Soil input data variation and clusters of similar soils

We defined 10 clusters of soils sharing similar soil properties (Section 2.4.1). However two of these clusters were associated with highly variable model outputs and were therefore additionally split into two groups (Section 2.4.2). First, shallow clay-rich soils were separated from other shallow soils and second, SOC-rich sandy soils were separated from the other sandy soils. This resulted in a total of 12 clusters (Table 2). The between-cluster variation accounted for 68% of the total variation in the soil data, while the within-cluster variation varied between 0.6 (cluster 8) and 5.4% (cluster 7). Sandy soils were found in clusters 2 to 7 and silt loam soils were found in clusters 8 and 9. The soils in clusters 10 and 11 were characterized by a shallow profile depth (< 0.8 m, Table 2).

The twelve soil clusters covered from 1 to 27% of the area of the NRW region at the finest resolution (Fig. 1) and included between 22 and 436 individual soils (Table 2). Clusters 10, 9 and 3 were the most dominant ones and covered 27, 19 and 15% of the NRW region respectively.

3.2.2. Identification of key functional soil types

The simulated yields were highest for sandy soils (clusters 3 to 7), intermediate for silt loam and clay loam soils (clusters 8 to 10; Fig. 4A)

and lowest for the shallow soils (clusters 11 and 12). The lower yields on shallow soils (< 0.8 m) were due to the limited root depth on these soils which induced more severe or more frequent water and/or N stresses. Drainage was slightly higher in the sandy soils than in the silt loam soils and significantly lower in clay and clay loam soils, mainly because more water was lost via surface runoff, especially for shallow clay soils. However, within-cluster variation in drainage was high for the clay soils (clusters 10 and 11; Fig. 4B). C mineralisation and N leaching (Fig. 4C & D) were highest in SOC-rich soils (> 7.7% in topsoil; clusters 1 and 2; Table 2), but the within-cluster variation was also high. C mineralisation was lower in shallow soils (clusters 11 and 12) as they contained less SOM in total, whereas N leaching was higher than for other non SOC-rich soils.

Key functional soil types were identified in five of the clusters and were classified into three distinct groups:

1. SOC-rich soils (> 7.7% in the topsoil or > 0.8% in the root zone) in clusters 1 and 2 were associated with the highest C mineralization and N leaching.
2. Clay soils (> 25%) in clusters 10 and 11 were associated with the lowest drainage.
3. Shallow soils (< 0.8 m) in clusters 11 and 12 were associated with the lowest yields, low C mineralisation and high N leaching.

These *key soils* and their specific characteristics are highlighted in Table 2. Fig. 4 shows in relative terms (%) the deviations in the average simulated variables associated with these *key soils* (i.e. D_{ks}; Section 2.4.2) from the overall mean. These three groups of soils contained 2.2% (i.e. 58 out of 2646 soils), 14.9% (393) and 4.6% (123) respectively of the total soil population of soils in the NRW region database (Table 2).

3.3. Soil spatial diversity and spatial pattern of key soils

The functional diversity of the 2646 individual soils of the NRW region was limited (H_s = 0.32; Eq. (1)), whereas the diversity of the soil clusters was rather high (H_c = 0.83). This is because many of the soils covered a very small part of the region, while a few others dominated, while the distribution of the relative coverage of the 12 clusters was

Table 2
Soil clusters based on input data of soils present in the NRW region at 1 km resolution. Between-cluster variation equalled 68.6% and within-cluster variation varied between 0.6 and 5.4%. N is the cluster number and Nb is the number of soils inside the cluster. Mean value ± standard deviation of input variables for three soil layers are presented; key soils associated with extreme values of at least one output are highlighted in bold (Section 2.4.2 and Fig. 4).

|	Soil type	SOC rich	Sandy SOC	Sandy					Silt loam		Clay	Shallow Clay	Shallow
Soil profile	N Clusters	1	2	3	4	5	6	7	8	9	10	11	12
	Nb. soils	36	22	432	48	301	298	298	436	312	342	51	72
Topsoil	Thick. (m)	0.3 +/- 0	0.3 +/- 0	0.3 +/- 0	0.3 +/- 0	0.3 +/- 0	0.3 +/- 0	0.3 +/- 0	0.3 +/- 0	0.3 +/- 0	0.3 +/- 0	0.3 +/- 0.1	0.3 +/- 0.1
	Clay (%)	11 +/- 7	8 +/- 9	5 +/- 3	10 +/- 4	5 +/- 2	18 +/- 7	12 +/- 6	19 +/- 4	23 +/- 7	**26 +/- 8**	**39 +/- 5**	20 +/- 6
	Sand (%)	60 +/- 21	73 +/- 28	83 +/- 10	62 +/- 16	82 +/- 10	42 +/- 14	57 +/- 13	16 +/- 7	19 +/- 8	20 +/- 9	19 +/- 8	24 +/- 18
	SOC (%)	**8.6 +/- 2.2**	2.6 +/- 0.5	2.3 +/- 0.4	2 +/- 0.2	2.2 +/- 0.4	2.4 +/- 0.4	2.1 +/- 0.2	2.1 +/- 0.3	2.1 +/- 0.3	2.1 +/- 0.2	2.1 +/- 0.3	2.2 +/- 0.2
Root zone	Thick. (m)	1.2 +/- 0	1.2 +/- 0	1.2 +/- 0.1	1.2 +/- 0	1.2 +/- 0.1	1.2 +/- 0	1.2 +/- 0	1.2 +/- 0	1.2 +/- 0	1.2 +/- 0	**0.3 +/- 0.1**	**0.4 +/- 0.2**
	Clay (%)	9 +/- 4	7 +/- 8	5 +/- 2	8 +/- 3	7 +/- 4	13 +/- 5	19 +/- 8	20 +/- 4	21 +/- 7	**33 +/- 7**	**37 +/- 3**	29 +/- 6
	Sand (%)	68 +/- 14	77 +/- 25	85 +/- 8	71 +/- 11	75 +/- 12	61 +/- 11	48 +/- 10	18 +/- 6	30 +/- 12	23 +/- 8	27 +/- 6	29 +/- 12
	SOC (%)	0.7 +/- 0.3	**1.0 +/- 0.2**	0.3 +/- 0.1	0.5 +/- 0.1	0.3 +/- 0.1	0.3 +/- 0.1	0.4 +/- 0.1	0.3 +/- 0.1	0.4 +/- 0.1	0.4 +/- 0.1	0.5 +/- 0.4	0.6 +/- 0.5
Subsoil	Thick. (m)	0.6 +/- 0.1	0.5 +/- 0.1	0.5 +/- 0.1	0.5 +/- 0.1	0.5 +/- 0.1	0.5 +/- 0.1	0.5 +/- 0.1	0.5 +/- 0.1	0.6 +/- 0.1	0.6 +/- 0.1		
	Clay (%)	9 +/- 5	7 +/- 9	4 +/- 2	5 +/- 2	16 +/- 10	6 +/- 3	30 +/- 9	24 +/- 9	7 +/- 4	**35 +/- 8**		
	Sand (%)	76 +/- 14	83 +/- 20	86 +/- 6	84 +/- 7	50 +/- 14	84 +/- 10	36 +/- 11	28 +/- 13	80 +/- 12	31 +/- 10		
	SOC (%)	0.1 +/- 0.2	1 +/- 0.2	0.1 +/- 0.1	0.5 +/- 0	0.2 +/- 0.1	0.1 +/- 0.1	0.2 +/- 0.1	0.2 +/- 0.1	0.2 +/- 0.1	0.2 +/- 0.1		
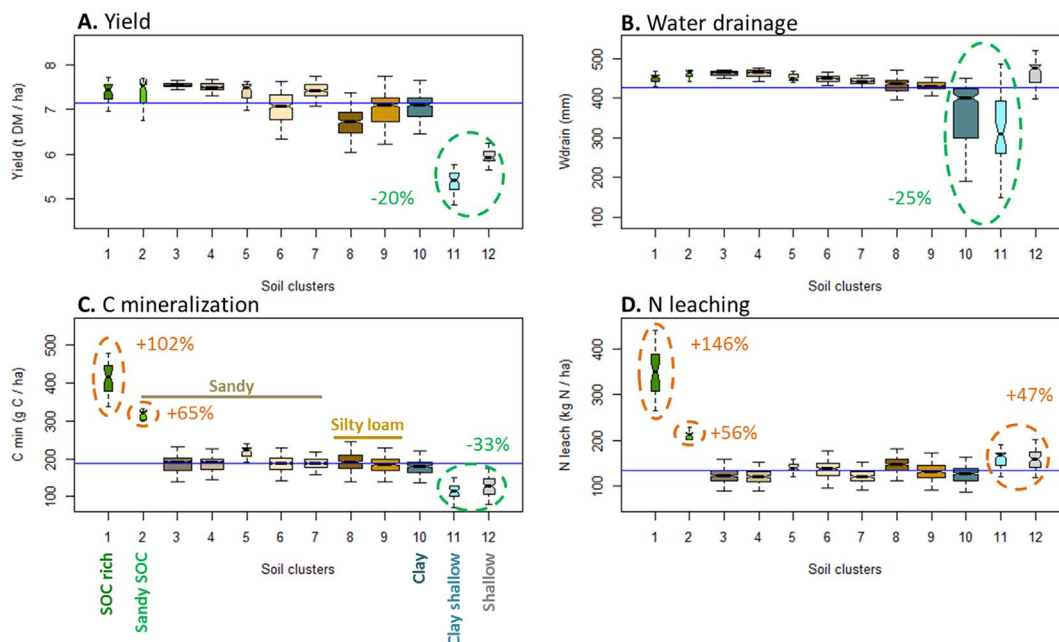
Fig. 4. Variation in simulated outputs within and between the 12 soil clusters. The horizontal lines show the average value for *all soils* present in the database ($S_{allsoils}$); dotted circles show the identified clusters containing '*key soils*' (Section 2.4.2). The relative boxplot widths are proportional to the number of soils included in each cluster (see also Table 1).

relatively more uniform.

The four sub-areas differed in the diversity and patterns of the soils present, but all the soil clusters were represented in all four sub-areas, with the exception of SOC-rich soils in area '*d*' (Table 3). The north-west area ('*a*') had the highest soil diversity while the south-east area ('*d*') had the lowest diversity (Table 3). This difference in diversity was more pronounced when considering the distribution of soil clusters. The diversity in soils and soil clusters was similar in sub-regions '*b*' and '*c*' (Table 3).

The *key soils* were unevenly distributed among the four sub-areas. Shallow soils were mostly present in the southern part of the region (sub-areas '*b*' and '*d*'), while SOC-rich soils were concentrated in the north-west ('*a*') and clay soils were dominant in the south-east ('*d*'). The distributions of *key soils* were spatially aggregated with the exception of SOC-rich soils in area '*b*' where they covered < 1%. The degree of spatial aggregation (*AI*, Section 2.5; Table 3) was highest for the shallow soils in area '*c*' and for the clay soils in '*d*', where the latter were also more abundant. Even though the aggregation index (*AI*) accounts

for differences in coverage, it was positively correlated to the soil coverage and to the uniformity index (H_{ks}).

3.4. DAE on the coverage of key soils

Fig. 5 illustrates the coverage of *key soils* in NRW at the different data resolutions. In relative terms, the coverage of SOC-rich soils was most affected by the data aggregation, with an increase of 130% at a resolution of 50 km and a decrease of 100% at a resolution of 100 km compared to the reference. However, they covered < 2% of the total area in the reference simulations. In absolute terms, the coverage of clay soils and shallow soils was affected more, varying by up to 4–5% which in relative terms represents changes of 17% and 50% respectively (Fig. 5). The sum of the partial areas that were converted between *key* and *non-key soils* (a_{ks} and a_0, checked colours in Fig. 5) as a result of data aggregation was highest for the clay soils in absolute terms (Fig. 5B), but highest for the SOC-rich and shallow soils in relative terms (Fig. 5A and C). Fig. 6 shows that the spatial patterns of *key*

Table 3
Diversity of soils and spatial pattern of key soils in the four sub-areas of NRW: normalized Shannon Index of diversity for soils (H_s) (as well as total number of soils into brackets), for soil clusters (H_c) and for key soils (H_{ks}, uniformity index) which varies from 0 (low values associated with dominance of one soil) to 1 (all soils tend to have equal coverage) as well as the spatial coverage (cov, %) and the normalized index of spatial aggregation of key soils versus *non-key soils AI* which varies between 0 (total dispersion) to 1 (maximal aggregation); refers to Section 2.5.

| All soils or *key soils*		West (a, b)				East (c, d)			
		Diversity [0–1]	Coverage [0–100]	Uniformity H_{ks} [0–1]	Aggregation *AI* [0–1]	Diversity [0–1]	Coverage [0–100]	Uniformity H_{ks} [0–1]	Aggregation *AI* [0–1]
North (a, c)	Soils H_s	0.88 (1001)				0.83 (705)			
	Clusters H_{ks}	0.85				0.76			
	SOC rich		05.9%	0.32	0.85		1.3%	0.10	0.64
	Clay		06.5%	0.35	0.81		22.1%	0.72	0.94
	Shallow		02.6%	0.17	0.59		04.6%	0.27	0.77
South (b, d)	Soils H_s	0.80 (694)				0.68 (435)			
	Clusters H_{ks}	0.79				0.47			
	SOC rich		0.6%	0.05	0.00		00.0%	0.00	–
	Clay		16.7%	0.70	0.93		71.7%	0.86	0.98
	Shallow		18.3%	0.69	0.91		09.7%	0.46	0.85
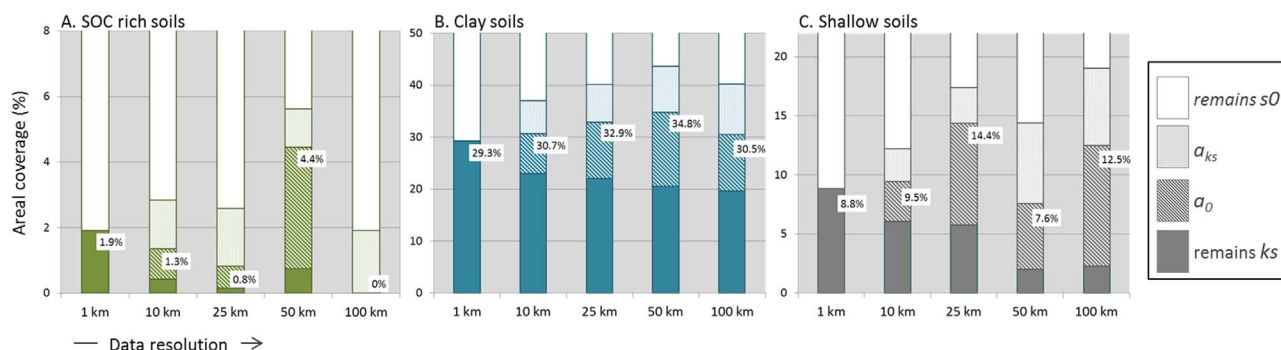
Fig. 5. Spatial coverage (% in labels) of the three key functional soil types ('*key soils*') at the different resolutions (full colour: area that remains *key soils* + gridded colour: area converted to *key soils* a$_0$ – because it was locally dominant), as well as the area where key soils were converted to non-key soils aks - because they were not locally dominant (rutted white).

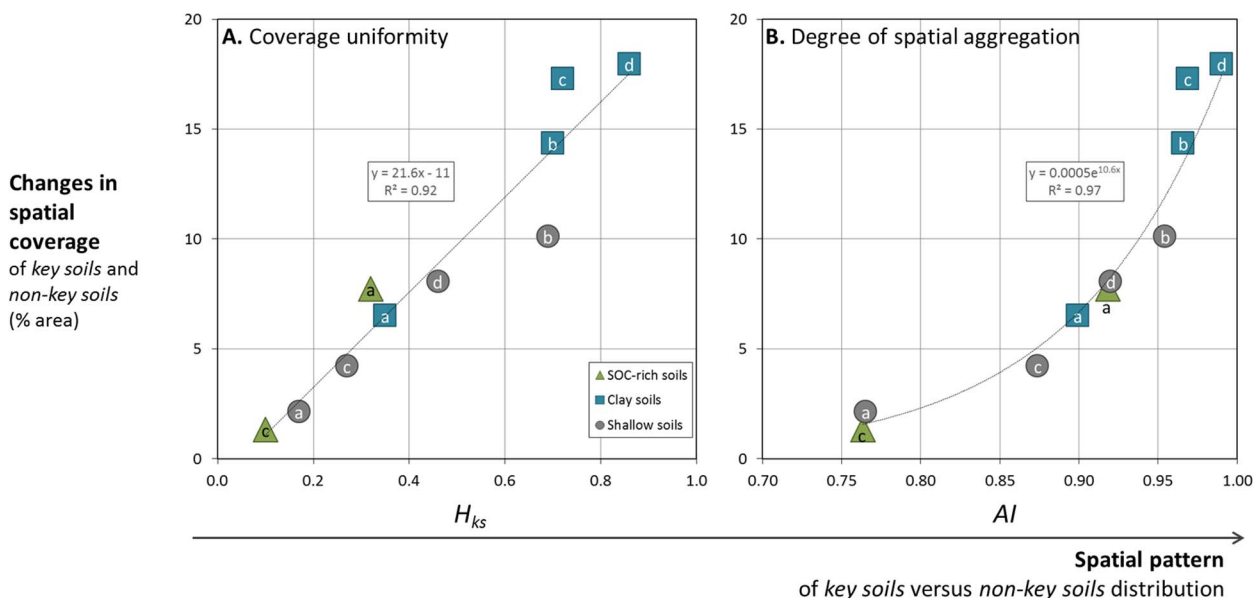**Fig. 6.** Effect of the spatial pattern of key soils versus non-key soils (coverage uniformity A. and degree of spatial aggregation B.) in the reference (i.e. at resolution 1 km) on the % area that was either converted from key soil to non-key soil or from non-key soil to key soil (i.e. sum of Σaks/ΣAks and Σa0/ΣA0); see also Fig. 3) when aggregating data from resolution 1 km to resolution 10 km. The labels 'a', 'b', 'c' and 'd' denote the four sub-areas of the NRW region (see also Fig. 1). The slope of the linear regression (A.) and the exponential coefficient (B.) were higher at resolutions 50 and 100 km than at resolutions 10 and 25 km. The linear regressions (A.) had coefficients R = 0.96; 0.94; 0.83 and 0.70 (P < 0.05) respectively for data aggregated at 10, 25, 50 and 100 km (data not shown for 25 to 100 km resolution). The respective values for the exponential regressions (B.) were R = 0.98; 0.84; 0.94 and 0.85 (P < 0.05).

and *non-key soils* in the reference map at 1 km resolution influenced the size of these converted partial areas, which we hypothesize to be mostly responsible for the DAE. The converted areas increased linearly with the uniformity index for the *key soils* (H_{ks}) and exponentially with the degree of spatial aggregation (*AI*).

3.5. DAE on model simulations

3.5.1. Regional bias and regional mean error

The regional bias caused by data aggregation was relatively small (*Bias*, Eq. (2) in 2.6) for all regional simulated variables. The absolute regional bias for yield and drainage simulated at 10, 25, 50 and 100 km resolutions was < 1.5% (Table 1) and was only slightly higher for C mineralisation and N leaching (< 6%). There was no clear trend in bias with respect to data resolution. This result is in line with previous modelling studies on crop yields in the same region (Angulo et al., 2014; Hoffmann et al., 2016a) and on water flow and N leaching in another region of Germany (Kersebaum and Wenkel, 1998). However, regional bias does not explicitly account for spatial differences within the region and compensating errors might occur (Zhao et al., 2015a). Thus, local discrepancies might still be large at the different resolutions. This highlights the importance of exploring the spatial patterns of the

output variables in regional-scale modelling (Steffens et al., 2015; Vereecken et al., 2016).

The DAE expressed as the regional mean error (*rRMSE*, % *in* 2.6 & Table 1a of A2) was larger for C mineralisation (15%) and N leaching (24%) than for yield and drainage (< 8%), at all resolutions (Table 1). This is in line with the results of multi-model comparison studies carried out for NRW with respect to SOC changes (Grosz et al., 2017) and N dynamics (Haas et al., 2015) compared with yield (Hoffmann et al., 2016a) and NPP (Kuhnert et al., 2016). Both C mineralisation and N leaching depend strongly on C stocks which varied widely among soils (4–73 kg m^{-2}), so coarse resolution soil data can lead to large biases in simulations of SOC (Zhang et al., 2016, 2016b). Only small differences in the DAE were found among the different resolutions, except for C mineralisation and N leaching, where it was much larger at a resolution of 50 km.

Expressing the regional bias as absolute values and considering all output variables and all resolutions together, the two DAE metrics were significantly correlated with each other. The average *rRMSE* calculated across the four resolutions (6.5%, 7.7%, 14.8%, 23.8% for yield, drainage, C mineralization and N leaching respectively) was exponentially correlated to the spatial variability of the variables (y = 3.6 e$^{0.07x}$; R^2 = 0.97; not shown) as previously shown by Zhao et al. (2015a) and Zhao et al. (2015b).
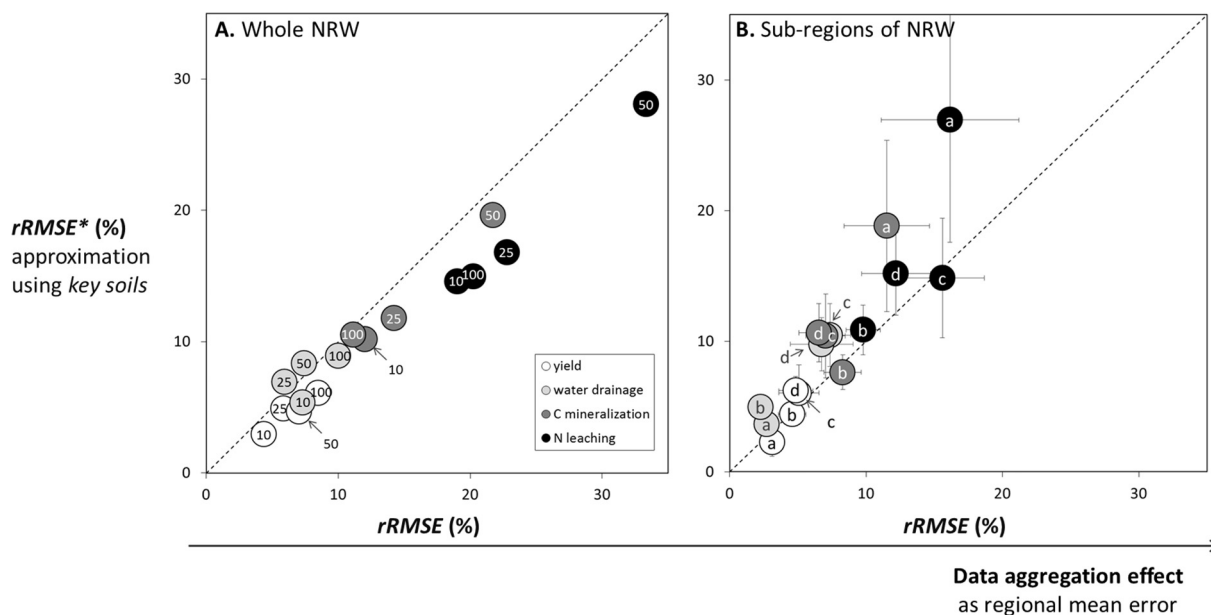
Fig. 7. Approximated DAE (rRMSE*_%) in NRW using *key soils* (2.7.3 and Table 1b in A2) against the true DAE (rRMSE_%; 2.7.2 and Table 1a in A2) calculated for the different simulated variables: (A.) individual values obtained at the different soil data resolutions (10, 25, 50 and 100 km) in the whole NRW region and (B.) mean values and standard deviation of the four different resolutions (10, 25, 50 and 100 km) in the four sub-areas of NRW (*a*, *b*, *c* and *d*; see also Fig. 1).

3.5.2. Approximation of the DAE based on key soils

Generally, the presence of certain soils ('*key soils*') in the NRW region or in sub-areas of the region was important for the DAE expressed as the regional mean error (*rRMSE*; Table 1a in Appendix 2). Considering all variables and resolutions together, the DAE was significantly correlated with its approximation *rRMSE** ($R^2 = 0.96$; $n = 16$). This was also the case when considering differences in DAE between the sub-areas of NRW and between variables (Fig. 7B; $R^2 = 0.78$ when considering all variables, sub-regions and resolutions together – $n = 64$). *rRMSE** captured slightly less of the difference in DAE between sub-areas than between variables and resolutions.

The approximation generally overestimated the DAE in the sub-areas, whereas it underestimated the DAE for the whole NRW region (slopes of the regressions are 1.4 and 0.8 respectively), especially in the case of N leaching. The overestimation for the sub-areas was especially high for C mineralisation and N leaching in the north-west area ('*a*'), which has the highest coverage of SOC-rich soils. In the whole region, the underestimation might be caused by differences within and between soil clusters of the *non-key soils* lost in the approximation of the DAE. The distributions of soils are more homogeneous in the sub-areas than in the whole region because of the spatial correlations of soil properties (see e.g. Meersmans et al., 2008). As a result, the variation in soil properties was less in the sub-areas, as were the differences between *key soils* and *non-key soils*. Consequently the calculated deviation (D_{ks}) was most likely overestimated in the sub-regions.

The DAE was higher at resolutions for which the change in the coverage of *key soils* compared with the reference simulations was greatest. For example, the DAE for N leaching and C mineralisation was highest at a resolution of 50 km where the spatial coverage of the SOC-rich soils was more than twice that in the reference (Figs. 5 & 7). Similarly, the change in the spatial coverage of shallow soils was greatest at a resolution of 100 km, intermediate at resolutions of 25 and 50 km and minimal at a resolution of 10 km (Fig. 5). This was mirrored by the variation in DAE for the yield in NRW (Fig. 7A).

3.5.3. Influence of the spatial pattern of key soils

The north-west area ('*a*') showed the highest DAE for C mineralization and N leaching, mainly as a result of the concentration of almost all SOC-rich soils in this area (Fig. 7B and Table 1). In contrast,

the DAE was lowest for yield and drainage in this area (Fig. 7B, Table 1) as a result of smaller values of the uniformity index for shallow soils and clay soils (H_{ks}; Table 3). The eastern area ('*c*' & '*d*') associated with a dominance of clay soils, had the highest DAE for drainage (Fig. 7B and Table 1).

The global differences in the DAE between sub-areas (i.e. considering all variables and resolutions together) did not correlate with differences in the diversity indexes for soils or soil clusters (Table 3). Moreover, the approximation of the DAE predicted the lowest DAE in area '*b*', but failed to predict the highest DAE found in sub-area '*c*'. Instead the approximation predicted the highest value in area '*a*', where the difference between the DAE and its approximation was greatest. One possible explanation for this mismatch between DAE and the soil spatial patterns is that differences between the sub-areas were too small. Furthermore, the two indicators used to describe spatial diversity (coverage and aggregation) were highly correlated, which made it difficult to isolate their individual effects on the DAE. Larger values of the DAE were obtained for sub-regions where *key soils* covered a larger part of the area.

The normalized Shannon diversity index (uniformity index) based on the relative distribution of *key soils* versus *non-key soils* may be a better indicator of the effects of soil spatial pattern, as it does not distinguish between which of these two groups is dominant (*key soils* or *non-key soils*). However, this hypothesis could not be properly tested as *key soils* were dominant only in the case of Clay soils in the sub-area '*d*', so that the difference between the coverage of *key soils* and the Shannon index was small. Nevertheless, the Shannon index was positively correlated with the partial areas converted between *key* and *non-key soils* when aggregating to coarser resolutions (Fig. 6A), which, in turn, influences the DAE.

3.6. Implications for regional crop modelling

Our analysis showed that a few soils generated large upscaling errors in model simulations when aggregating soil data based on the "dominant soil concept" (selecting the soil with largest areal coverage). The pedodiversity of a region is generally regarded as an important factor influencing the effect of soil data aggregation on model simulations (Jégo et al., 2015). Our results suggest that the pedodiversity per

se (Costantini and L'Abate, 2016) may be less critical than its effect on soil-crop processes (the *functional diversity* of soils) for estimating the DAE. Similarly, decision-makers are more interested in the accuracy of maps of land quality which show the response of environmental variables to soil variability than of the underlying soil maps themselves (Hennings, 2002; Takoutsing et al., 2017). The *functional diversity* of soils was assessed here by a model sensitivity analysis to different clusters of soils (sharing similar soil properties) present in the region and by isolating clusters of soils ('*key functional soil types*') that generated extreme outputs. The identification of critical soil properties depends on environmental conditions such as climate and management practices (Folberth et al., 2016). Therefore, even though the *key functional soil types* identified here are likely to be relevant elsewhere, they are at least partly specific to the regional context (pedo-climatic conditions and cropping systems) and to the modelling approach applied (model structure and parameterisation). However, to reduce the potential aggregation errors and also the computational time when applying soil-crop models at regional scales, the methodology developed here for screening functional groups of soils and sensitive soils (in relation to the modelling approach chosen) should prove useful and could be applied in other environmental contexts and with other models.

4. Conclusions

Our results suggest that the grid resolution in regional-scale simulations of cropping systems should be adapted to soil variability and to the target output variables. A screening method of the spatial distribution of critical functional soil types (for which the target variable is sensitive) should help in designing an optimal variable mesh: where do we need a fine resolution and where can we use a coarse resolution? For example, areas without these critical soils may be simulated at much coarser scales than areas where they are present. If these soils cover only a very small fraction of the region, their influence on simulations at the regional scale will most likely be negligible, although they may have significant impacts locally. Our results further suggest that the influence of the spatial pattern of these key (e.g. spatial aggregation) soils should be considered when designing the grid resolution in regional-scale simulations. However, relevant spatial pattern indicators that can be linked to DAE still needs further development.

Appendix 1. Model options and parameterisation

1.1. Crop growth

Plant growth was modelled using a radiation use efficiency approach combined with response functions for unfavourable temperature, N and water conditions, including effects of simulated soil temperature and water and nutrient availability as influenced by soil properties and weather. The allocation parameters for winter wheat growth were taken from Conrad and Fohrer (2009c). Thereafter, a few parameters were further adjusted to fit the model predictions to target (observed) values for a regional time averaged actual annual yield (DM) of 7.2 t ha^{-1} (average for 1982–2010 in the NRW region), a harvest date on the August 1st and an assumed yield gap of ca. 15–20%. Firstly, sowing and emergence were set at the fixed days given by Hoffmann et al. (2015). Secondly, the temperature sum needed for achieving maturity was adjusted to fit the regional harvest date using a regional averaged climate. Thereafter, the radiation use efficiency and the parameters governing leaf area index development were fitted to achieve a regional and period averaged potential yield (DM) of 9 t/ha (estimated as actual yield divided by yield gap). Before adjusting simulated actual yield, the maximum root depth was set to 1.5 m according to Hoffmann et al. (2015). Thereafter, parameters regulating crop growth as limited by water (CritThresholdDry, Table 1 - Plant Growth) and N uptake (Flexibility Degree) by the roots were manually fitted to reach the regional actual yield.

1.2. Soil hydraulic functions and bottom boundary conditions

The Brooks-Corey-Mualem model was used to describe soil water retention and hydraulic conductivity (Brooks and Corey, 1964; Mualem, 1976). The model parameters were estimated with the Rawls and Brakensiek pedotransfer functions (Rawls and Brakensiek, 1989) from the clay and sand contents, total soil porosity and water content at wilting point from the soil database. A coupling between the NRW-soil database and the CoupModel was developed to facilitate the multiple model-runs required to cover the large number of grid cells over the region. The water flow from the bottom layer was calculated from the unsaturated conductivity of the bottom layer, assuming a unit hydraulic gradient (i.e. gravity-driven flow). Field drainage systems, macropores flows and groundwater flows were not considered.

1.3. Soil C and N

The initial C and N content of the soil organic pools (litter and humus) were set according to soil C content values in the NRW database and based on the assumption that 1.2% of the total C consisted of litter and the remaining part was humus and that the C-N ratios of litter and humus were 25 and 10, respectively (giving a total soil C-N ratio of 10.5). The rate coefficient for the decomposition of soil organic matter and the potential denitrification rate were fixed to a unique value for all soils, since no data on variation among soils were available. The decomposition rate was adjusted manually to 6×10^{-5} d^{-1} to obtain 'reasonable' average C losses when using the dominant soil of the region (original value was 7 10^{-5} d^{-1}; Johnsson et al., 1987). The nitrification specific rate was taken from Johnsson et al. (1987) and the denitrification potential was adjusted from Conrad and Fohrer (2009b).

Table Appendix 1
Model options and parameters changed from default values used in this study.

	Model options and parameters (O, P)	References
Climate data – plant interactions		
P Latitude	**50.3269**	Input data
P Reference height	**1.5 m** – *for air temperature, humidity and wind speed*	Input data
P Albedo dry	**23%** – *albedo of a dry soil*	Gustafsson et al., 2004
P Ra Increase With LAI	**28.9016 s m^{-1}** – *contribution of LAI to the total aerodynamic resistance from measurement height*	Gustafsson et al., 2004
Plant growth		
O Growth	**Radiation use efficiency**	–
O Plant stress	**All multiplicative** – *response functions for water, heat and nitrogen*	
O Plant type	**Explicit big leaves** – *separation between transpiration and evaporation*	
O Initial plant conditions	**As nitrogen**	–
P Specific Leaf Area (SLA)	**0,038 g m^{-2}** – *parameter to convert leaf C mass into leaf area*	Adjusted
O Plant development	**Start = f (Temp Sum)** – *start of growth stages is a function of temperature sums.*	
O Harvest day	**Simulated** (when maturity occurs)	–
O Root input	**Simulated**	–
O Root distribution with depth	**Exponential**	–
P Root lowest depth	**− 1.5 m** (*or the depth of soil if soil < 1.5 m*)	Input data
P Flexibility degree	**0.5** (no unit)	Adjusted
P CritThreshold dry	**200 cm** – *critical pressure head for reduction of potential water uptake*	Adjusted
Soil profile and hydraulic properties		
P Number of model layers & thickness	**18** (soils > 1.2 m: **0.05 m × 2, 0.1 m × 8, 0.2 m × 4, 0.3 m × 2**) or **9** (soils ≤ 0.8 m: **0.05 m × 2, 0.1 m × 7**)	–
O Hydraulic conductivity	**Mualem equation**	–
O Soil water flow input	**Simulated** – *soil water flow is simulated*	–
O Soil water input	**Simulated** – *soil water content is simulated*	–
O Hydraulic functions	**Brooks & Corey**	–
O Soil infil input	**Simulated** – *soil water infiltration is simulated*	–
O Soil drainage input	**Not used** – *no drainage is considered*	–
O Deep percolation input	**Simulated** – *values are simulated*	–
O Ground water flow	**Off**	–
O Lower boundary unsaturated	**Unit gradient gravitational flow**	–
O Brooks-Corey	**Rawls & Brakensiek** pedotransfert function	–
Soil organic matter and nutrients		
O Soil organic processes	**Separated litter and humus pool**	–
O Initial carbon conditions	**As nitrogen and carbon**	–
O Nitrogen and carbon	**Dynamic interaction with abiotics**	–
P Denitrification potential rate	**0.27 g m^2 d^{-1}**	Adjusted from Conrad and Fohrer, 2009b
P Specific nitrification rate	**0.2 d^{-1}**	Johnsson et al., 1987
P Rate coefficient for humus decomposition	**6 ∗ 10^{-5} d^{-1}**	Adjusted from Johnsson et al., 1987
P Initial mineral-N concentration	**Nitrate-N = 10 mg l^{-1} & ammonium-N = 10 mg l^{-1}**	*Default values*
P Initial organic C & N	*Calculated by the CoupModel based on soil inputs data*	Input data
Management practices		(Hoffmann et al., 2015)
P Sowing DayNo	**274** (Day Of Year; DOY)	Input data
P C seed	**10 g m^{-2}** – *carbon content in seeds*	Input data
P Emergence DayNo	**280** (DOY)	Input data
P N fertilization	**3** – *number of fertilisation events*	Input data
P Fert DayNo	**60, 105, 152** (DOY)	Input data
P N Fert rate	**13, 5.2, 2.6 g m^{-2} d^{-1}**	Input data
O Irrigation	**Off**	Input data
P Ploughing DayNo	**244**	Input data

Appendix 2. Metrics for soil Data Aggregation Effects (DAE)

Table 1a
Calculation of the DAE expressed as the regional mean error presented in Section 2.6.

DAE as regional mean error			
$rRMSE$ (%)	$rRMSE = (RMSE/\overline{S_r})\,100$	(3)	• $\overline{S_r}$ is the regional average of the reference simulations (at resolution 1 km)
	$rRMSE = \left(\sqrt{\sum_{i=1}^{N_c} (A_{ci}\,(S_{ci} - \overline{S_{r,c}})^2)/A_T} \, / \overline{S_r} \right) 100$	(4)	• Eq. (4) is the developed equation for $rRMSE$ and the different terms are presented below
$RMSE$ Area-weighted	$RMSE = \sqrt{\sum_{i=1}^{N_c} (A_{ci}\,e_i^2)/A_T}$	(5)	• N_c is the number of coarse grids • A_{ci} is the area of the coarse grid i • A_T is the total area of the NRW region
e_i Error associated to one coarse grid i	$e_i = S_{ci} - \overline{S_{r,c}}$	(6)	• S_{ci} is the true simulated variable
$\overline{S_{r,c}}$ Mean simulated variable at resolution 1 km within the coarse grid i	$\overline{S_{r,c}} = \sum_{j=1}^{n_1} S_{rij}/n_1$	(7)	• S_{rij} is the simulated variable in one grid cell j of the reference resolution (1 km) • n_1 is the number of 1 km² grid cells inside the coarse grid i

Table 1b
Approximation of the DAE based on *key soils*, expressed as the regional mean error and presented in Section 2.7.

Approximated regional mean error			
$rRMSE^*$ (%)	$rRMSE^* = (RMSE^*/\overline{S_r})\,100$	(3)	• $\overline{S_r}$ is the regional average of the reference simulations (at resolution 1 km)
	$rRMSE^* = \left(\sqrt{\left(\sum_{i=1}^{N_{ks}} (a_{0_i}\,D_{ks})^2/A_{ksi} + \sum_{j=1}^{N_0} (a_{ksj}\,D_{ks})^2/A_{0j} \right)/A_T} \, /\overline{S_r} \right) 100$	(8)	• Eq. (8) is the developed equation for $rRMSE^*$ and the different terms are presented below
$RMSE^*$ Area-weighted	$RMSE^* = \sqrt{\left(\sum_{i=1}^{N_{ks}} A_{ksi}\,e1_i^2 + \sum_{j=1}^{N_0} A_{0j}\,e2_j^2 \right)/A_T}$	(9)	• N_{ks} and N_0 are the number of coarse grids in the region that are *key soils* and *non-key soils* respectively • A_{ksi} and A_{0j} are the respective area covered by *key soils* and *non-key soils* within the coarse grid i and j respectively • A_T is the total area of the NRW region
$e1_i$ and **$e2_j$** Error associated to one coarse grid i or j	$e1 = S_{ci}^* - \overline{S_{r,c}^*} = (\overline{S_{allsoils}} + D_{ks}) - \overline{S_{r,c}^*} = D_{ks}\,a_0/A_{ks}$ $e2 = S_{cj}^* - \overline{S_{r,c}^*} = \overline{S_{allsoils}} - \overline{S_{r,c}^*} = -D_{ks}\,a_{ks}/A_0$	(10a) (10b)	• S_{ci} is approximated by $(\overline{S_{allsoils}} + D_{ks})$ in the case of **e1** (i.e. coarse grid is a *key soil*) and S_{cj} by $\overline{S_{allsoils}}$ in case of **e2** (i.e. coarse grid is *non-key soil*) • $\overline{S_{allsoils}}$ is the average of the simulations with *all soils* from the database and D_{ks} is the deviation in simulated variable between the *key soils* and the average of *all soils* (2.6.2) • a_0 and a_{ks} are the number of 1 km² grid cells within the coarse grid that are *non-key soils* and *key soils* respectively
$\overline{S_{r,c}^*}$ Mean simulated variable at resolution 1 km within the coarse grid i or j	$\overline{S_{r,c}^*} = (a_0/A_c)\,\overline{S_{allsoils}} + (a_{ks}/A_c)\,(\overline{S_{allsoils}} + D_{ks})$ $= \overline{S_{allsoils}} + (a_{ks}/A_c)\,D_{ks}$	(11)	• $A_c = a_{ks} + a_0$ is the area of the coarse grid i or j

References

Angulo, C., Rötter, R., Trnka, M., Pirttioja, N., Gaiser, T., Hlavinka, P., Ewert, F., 2013. Characteristic "fingerprints" of crop model responses to weather input data at different spatial resolutions. Eur. J. Agron. 49, 104–114. http://dx.doi.org/10.1016/j. eja.2013.04.003.

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. 61, 35–48. http://dx.doi.org/10.1016/j.eja.2014.07.005.

Balkovič, J., van der Velde, M., Schmid, E., Skalský, R., Khabarov, N., Obersteiner, M., Stürmer, B., Xiong, W., 2013. Pan-European crop modelling with EPIC:

implementation, up-scaling and regional crop yield validation. Agric. Syst. 120, 61–75. http://dx.doi.org/10.1016/j.agsy.2013.05.008.

Brooks, R.H., Corey, A.T., 1964. Hydraulic properties of porous media and their relation to drainage design. Trans. ASABE 26–0028.

Cécillon, L., Barré, P., 2015. Soil functional types: surveying the biophysical dimensions of soil security. In: EGU Gen. Assem. 2015, Held 12-17 April. 2015 Vienna, Austria. ID. 7068 17.

Cliff, A.D., Ord, J.K., 1973. Spatial autocorrelation. Pion, London.

Confalonieri, R., Acutis, M., Bellocchi, G., Donatelli, M., 2009. Multi-metric evaluation of the models WARM, CropSyst, and WOFOST for rice. Ecol. Model. 220, 1395–1410. http://dx.doi.org/10.1016/j.ecolmodel.2009.02.017.

Conrad, Y., Fohrer, N., 2009a. Modelling of nitrogen leaching under a complex winter wheat and red clover crop rotation in a drained agricultural field. Phys. Chem. Earth 34, 530–540. http://dx.doi.org/10.1016/j.pce.2008.08.003.

Conrad, Y., Fohrer, N., 2009b. Application of the Bayesian calibration methodology for the parameter estimation in CoupModel. Adv. Geosci. 21, 13–24. http://dx.doi.org/10.5194/adgeo-21-13-2009.

Conrad, Y., Fohrer, N., 2009c. A test of CoupModel for assessing the nitrogen leaching in grassland systems with two different fertilization levels. J. Plant Nutr. Soil Sci. 172, 745–756. http://dx.doi.org/10.1002/jpln.200800264.

Costantini, E.A.C., L'Abate, G., 2016. Beyond the concept of dominant soil: preserving pedodiversity in upscaling soil maps. Geoderma 271, 243–253. http://dx.doi.org/10.1016/j.geoderma.2015.11.024.

Coucheney, E., Buis, S., Launay, M., Constantin, J., Mary, B., García de Cortázar-Atauri, I., Ripoche, D., Beaudoin, N., Ruget, F., Andrianarisoa, K.S., Le Bas, C., Justes, E., Léonard, J., 2015. Accuracy, robustness and behavior of the STICS soil–crop model for plant, water and nitrogen outputs: evaluation over a wide range of agro-environmental conditions in France. Environ. Model. Softw. 64, 177–190. http://dx.doi.org/10.1016/j.envsoft.2014.11.024.

De Gryze, S., Lee, J., Ogle, S., Paustian, K., Six, J., 2011. Assessing the potential for greenhouse gas mitigation in intensively managed annual cropping systems at the regional scale. Agric. Ecosyst. Environ. 144, 150–158. http://dx.doi.org/10.1016/j.agee.2011.05.023.

DWD, 2014. REGNIE Gridded Data of Daily Precipitation. German Meteorological Service (DWD). http://werdis.dwd.de/werdis_en/WebWerdis_start.do, Accessed date: 2 August 2014.

Eckelmann, W., Sponagel, H., Grottenthaler, W., Hartmann, K.-J., Hartwich, R., Janetzko, P., Joisten, H., Kühn, D., Sabel, K.-J., Traidl, R., 2005. Ad hoc Arbeitsgruppe Boden: Bodenkundliche Kartieranleitung, 5th ed. Federal Institute for Geosciences and Natural Ressources in cooperation with the Federal Geological Services, Hannover.

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. Agric. Ecosyst. Environ. 142, 6–17. http://dx.doi.org/10.1016/j.agee.2011.05.016.

Ewert, F., Rötter, R.P., Bindi, M., Webber, H., Trnka, M., Kersebaum, K.C., Olesen, J.E., van Ittersum, M.K., Janssen, S., Rivington, M., Semenov, M.A., Wallach, D., Porter, J.R., Stewart, D., Verhagen, J., Gaiser, T., Palosuo, T., Tao, F., Nendel, C., Roggero, P.P., Bartošová, L., Asseng, S., 2015. Crop modelling for integrated assessment of risk to food production from climate change. Environ. Model. Softw. 72, 287–303. http://dx.doi.org/10.1016/j.envsoft.2014.12.003.

Folberth, C., Yang, H., Wang, X., 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. Model. 235–236, 8–18. http://dx.doi.org/10.1016/j.ecolmodel.2012.03.035.

Folberth, C., Skalský, R., Moltchanova, E., Balkovič, J., Azevedo, L.B., Obersteiner, M., van der Velde, M., 2016. Uncertainty in soil data can outweigh climate impact signals in global crop yield simulations. Nat. Commun. 7, 11872. http://dx.doi.org/10.1038/ncomms11872.

Gaiser, T., Abdel-Razek, M., Bakara, H., 2009. Modeling carbon sequestration under zero-tillage at the regional scale. II. The influence of crop rotation and soil type. Ecol. Model. 220, 3372–3379. http://dx.doi.org/10.1016/j.ecolmodel.2009.08.001.

Geological Service NRW, 2004. Bodenkarte 1: 50.000 (BK50). Karte der schutzwürdigen Böden. Bearbeitungsmaßstab 1:50.000. CD-ROM. Geologischer Dienst Nordrhein-Westfalen. Krefeld. http://www.gd.nrw.de/g_bk50d.htm.

Gervois, S., Ciais, P., de Noblet-Ducoudré, N., Brisson, N., Vuichard, N., Viovy, N., 2008. Carbon and water balance of European croplands throughout the 20th century. Glob. Biogeochem. Cycles 22http://dx.doi.org/10.1029/2007GB003018. (n/a–n/a).

Giltrap, D.L., Li, C., Saggar, S., 2010. DNDC: a process-based model of greenhouse gas fluxes from agricultural soils. Agric. Ecosyst. Environ. 136, 292–300. http://dx.doi.org/10.1016/j.agee.2009.06.014.

Grassini, P., van Bussel, L.G.J., Van Wart, J., Wolf, J., Claessens, L., Yang, H., Boogaard, H., de Groot, H., van Ittersum, M.K., Cassman, K.G., 2015. How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis. F. Crop. Res. 177, 49–63. http://dx.doi.org/10.1016/j.fcr.2015.03.004.

Grosz, B., Dechow, R., Gebbert, S., Hoffmann, H., Zhao, G., Constantin, J., Raynal, H., Wallach, D., Coucheney, E., Lewan, E., Eckersten, H., Specka, X., Kersebaum, K.-C., Nendel, C., Kuhnert, M., Yeluripati, J., Haas, E., Teixeira, E., Bindi, M., Trombi, G., Moriondo, M., Doro, L., Roggero, P.P., Zhao, Z., Wang, E., Tao, F., Rötter, R., Kassie, B., Cammarano, D., Asseng, S., Weihermüller, L., Siebert, S., Gaiser, T., Ewert, F., 2017. The implication of input data aggregation on up-scaling soil organic carbon changes. Environ. Model. Softw. 96, 361–377. http://dx.doi.org/10.1016/j.envsoft.2017.06.046.

Gustafsson, D., Lewan, E., Jansson, P.-E., 2004. Modeling water and heat balance of the boreal landscape—comparison of Forest and arable land in Scandinavia. J. Appl. Meteorol. 43, 1750–1767. http://dx.doi.org/10.1175/JAM2163.1.

Haas, E., Klatt, S., Kiese, R., Hoffmann, H., Zhao, G., Specka, X., Nendel, C., Sosa, C.,

Lewan, E., Eckersten, H., Yeluripati, J., Kuhnert, M., Tao, F., Rötter, P., Constantin, J., Raynal, H., Wallach, D., Teixeira, E., Grosz, B., Bach, M., Doro, L., Pier, P., Zhao, Z., Wang, E., Trombi, G., Bindi, M., Moriondo, M., Cammarano, D., Asseng, S., 2015. Responses of soil N2O emissions and nitrate leaching on climate input data aggregation: a biogeochemistry model ensemble study Overview 1 Scaling exercise (done by Holger, Matthias and Balázs). In: MACSUR Conference 2015 - Integrated Climate Risk Assessment in Agriculture & FoodAssessment in Agriculture & Food. FACCE MACSUR Reports 2, pp. D-C5.20.

Hartigan, J.A., Wong, M.A., 1979. A K-means clustering algorithm. Appl. Stat. 28, 100–108.

He, H.S., Dezonia, B.E., Mladenoff, D.J., 2000. An Aggregation Index (AI) to Quantify Spatial Patterns of Landscapes. pp. 591–601.

Hennings, V., 2002. Accuracy of coarse-scale land quality maps as a function of the up-scaling procedure used for soil data. Geoderma 107, 177–196. http://dx.doi.org/10.1016/S0016-7061(01)00148-3.

Hoffmann, H., Zhao, G., van Bussel, L., Enders, A., Specka, X., Sosa, C., Yeluripati, J., Tao, F., Constantin, J., Raynal, H., Teixeira, E., Grosz, B., Doro, L., Zhao, Z., Wang, E., Nendel, C., Kersebaum, K., Haas, E., Kiese, R., Klatt, S., Eckersten, H., Vanuytrecht, E., Kuhnert, M., Lewan, E., Rötter, R., Roggero, P., Wallach, D., Cammarano, D., Asseng, S., Krauss, G., Siebert, S., Gaiser, T., Ewert, F., 2015. Variability of effects of spatial climate data aggregation on regional yield simulation by crop models. Clim. Res. 65, 53–69. http://dx.doi.org/10.3354/cr01326.

Hoffmann, H., Zhao, G., Asseng, S., Bindi, M., Biernath, C., Constantin, J., Coucheney, E., Dechow, R., Doro, L., Eckersten, H., Gaiser, T., Grosz, B., Heinlein, F., Kassie, B.T., Kersebaum, K.-C., Klein, C., Kuhnert, M., Lewan, E., Moriondo, M., Nendel, C., Priesack, E., Raynal, H., Roggero, P.P., Rötter, R.P., Siebert, S., Specka, X., Tao, F., Teixeira, E., Trombi, G., Wallach, D., Weihermüller, L., Yeluripati, J., Ewert, F., 2016a. Impact of spatial soil and climate input data aggregation on regional yield simulations. PLoS One 11, e0151782. http://dx.doi.org/10.1371/journal.pone.0151782.

Hoffmann H, Enders A, Siebert S, Gaiser T, Ewert F. 2016b. Climate and soil input data aggregation effects in crop models. doi: https://doi.org/10.7910/DVN/C0J5BB, (Harvard Dataverse), V3. https://dataverse.harvard.edu/dataset.xhtml?persistentId=http://dx.doi.org/10.7910/DVN/C0J5BB

Jansson, P.-E., 2012. CoupModel: model use, calibration, and validation. Trans. ASABE 55, 1337–1346. http://dx.doi.org/10.13031/2013.42245.

Jansson, P.E., Karlberg, L., 2013. Coupled Heat and Mass Transfer Model for Soilplant - Atmosphere System. (WWW Document).

Jansson, P.-E., Moon, D.S., 2001. A coupled model of water, heat and mass transfer using object orientation to improve flexibility and functionality. Environ. Model. Softw. 16, 37–46. http://dx.doi.org/10.1016/S1364-8152(00)00062-1.

Jégo, G., Pattey, E., Mesbah, S.M., Liu, J., Duchesne, I., 2015. Impact of the spatial re-solution of climatic data and soil physical properties on regional corn yield predictions using the STICS crop model. Int. J. Appl. Earth Obs. Geoinf. 41, 11–22. http://dx.doi.org/10.1016/j.jag.2015.04.013.

Jeuffroy, M.-H., Casadebaig, P., Debaeke, P., Loyce, C., Meynard, J.-M., 2014. Agronomic model uses to predict cultivar performance in various environments and cropping systems. A review. Agron. Sustain. Dev. 34, 121–137. http://dx.doi.org/10.1007/s13593-013-0170-9.

Johnsson, H., Bergstrom, L., Jansson, P.-E., Paustian, K., 1987. Simulated Nitrogen Dynamics and Losses in a Layered Agricultural Soil. Ecosyst. Environ. 18. Elsevier Sci. Publ. B.V, pp. 333–356.

Kersebaum, K.C., Wenkel, K.-O., 1998. Modelling water and nitrogen dynamics at three different spatial scales – influence of different data aggregation levels on simulation results. Nutr. Cycl. Agroecosyst. 50, 313–319. http://dx.doi.org/10.1023/A:1009721218584.

Kravchenko, A.N., Bullock, D.G., 2000. Correlation of corn and soybean grain yield with topography and soil properties. Agron. J. 92, 75. http://dx.doi.org/10.1007/s100870050010.

Kuhnert, M., Yeluripati, J., Smith, P., Hoffmann, H., van Oijen, M., Constantin, J., Coucheney, E., Dechow, R., Eckersten, H., Gaiser, T., Grosz, B., Haas, E., Kersebaum, K.-C., Kiese, R., Klatt, S., Lewan, E., Nendel, C., Raynal, H., Sosa, C., Specka, X., Teixeira, E., Wang, E., Weihermüller, L., Zhao, Q., Zhao, Z., Ogle, S., Ewert, F., 2016. Impact analysis of climate data aggregation at different spatial scales on simulated net primary productivity for croplands. Eur. J. Agron. http://dx.doi.org/10.1016/j.eja.2016.06.005.

LANUV, 2014. North Rhine-Westphalia State Agency for Nature, Environment and Consumer Protection. Unit 33, soil conservation, Recklinghausen, Germany. http://www.lanuv.nrw.de, Accessed date: 25 January 2015.

Lenz-Wiedemann, V.I.S., Klar, C.W., Schneider, K., 2010. Development and test of a crop growth model for application within a global change decision support system. Ecol. Model. 221, 314–329. http://dx.doi.org/10.1016/j.ecolmodel.2009.10.014.

Meersmans, J., De Ridder, F., Canters, F., De Baets, S., Van Molle, M., 2008. A multiple regression approach to assess the spatial distribution of Soil Organic Carbon (SOC) at the regional scale (Flanders, Belgium). Geoderma 143, 1–13. http://dx.doi.org/10.1016/j.geoderma.2007.08.025.

Merino, A., Pérez-Batallón, P., Macías, F., 2004. Responses of soil organic matter and greenhouse gas fluxes to soil management and land use changes in a humid temperate region of southern Europe. Soil Biol. Biochem. 36, 917–925. http://dx.doi.org/10.1016/j.soilbio.2004.02.006.

Mualem, Y., 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. Water Resour. Res. 12, 513–522. http://dx.doi.org/10.1029/WR012i003p00513.

Ng, H.Y.F., Drury, C.F., Serem, V.K., Tan, C.S., Gaynor, J.D., 2000. Modeling and testing of the effect of tillage, cropping and water management practices on nitrate leaching in clay loam soil. Agric. Water Manag. 43, 111–131. http://dx.doi.org/10.1016/

S0378-3774(99)00050-5.

Olesen, J., Bøcher, P., Jensen, T., 2000. Comparison of scales of climate and soil data for aggregating simulated yields of winter wheat in Denmark. Agric. Ecosyst. Environ. 82, 213–228. http://dx.doi.org/10.1016/S0167-8809(00)00227-9.

Ramezani, H., 2012. A note on the normalized definition of Shannon's diversity index in landscape pattern analysis. Environ. Nat. Resour. Res. 2, 54–60. http://dx.doi.org/10.5539/enrr.v2n4p54.

Rawls, W.J., Brakensiek, D.L., 1989. Estimation of soil water retention and hydraulic properties. In: Unsaturated Flow in Hydrologic Modeling. Springer Netherlands, Dordrecht, pp. 275–300. http://dx.doi.org/10.1007/978-94-009-2352-2_10.

Riffaldi, R., Saviozzi, A., Levi-Minzi, R., 1996. Carbon mineralization kinetics as influenced by soil properties. Biol. Fertil. Soils 22, 293–298. http://dx.doi.org/10.1007/BF00334572.

Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A.C., Müller, C., Arneth, A., Boote, K.J., Folberth, C., Glotter, M., Khabarov, N., Neumann, K., Piontek, F., Pugh, T.A.M., Schmid, E., Stehfest, E., Yang, H., Jones, J.W., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. Proc. Natl. Acad. Sci. 111, 3268–3273. http://dx.doi.org/10.1073/pnas.1222463110.

Schaap, M., Leij, F.J., 1998. Database-related accuracy and uncertainty of pedotransfer functions. Soil Sci. 163, 765–779.

Siebert, S., Ewert, F., 2012. Spatio-temporal patterns of phenological development in Germany in relation to temperature and day length. Agric. For. Meteorol. 152, 44–57. http://dx.doi.org/10.1016/j.agrformet.2011.08.007.

Steffens, K., Jarvis, N., Lewan, E., Lindström, B., Kreuger, J., Kjellström, E., Moeys, J., 2015. Direct and indirect effects of climate change on herbicide leaching - a regional scale assessment in Sweden. Sci. Total Environ. 514, 239–249. http://dx.doi.org/10.1016/j.scitotenv.2014.12.049.

Takoutsing, B., Rodríguez Martín, J.A., Weber, J.C., Shepherd, K., Sila, A., Tondoh, J., 2017. Landscape approach to assess key soil functional properties in the highlands of Cameroon: Repercussions of spatial relationships for land management interventions. J. Geochem. Explor. http://dx.doi.org/10.1016/j.gexplo.2017.03.014.

Thorndike, R.L., 1953. Who belongs in the family? Psychometrika 18, 267–276.

Tilman, D., Balzer, C., Hill, J., Befort, B.L., 2011. Global food demand and the sustainable intensification of agriculture. Proc. Natl. Acad. Sci. 108, 20260–20264. http://dx.doi.org/10.1073/pnas.1116437108.

Timlin, D.J., Pachepsky, Y., Snyder, V.a., Bryant, R.B., 1998. Spatial and temporal variability of corn grain yield on a hillslope. Soil Sci. Soc. Am. J. 62, 764. http://dx.doi.org/10.2136/sssaj1998.03615995006200030032x.

Tubiello, F.N., Donatelli, M., Rosenzweig, C., Stockle, C.O., 2000. Effects of climate change and elevated CO2 on cropping systems: model predictions at two Italian locations. Eur. J. Agron. 13, 179–189. http://dx.doi.org/10.1016/S1161-0301(00)00073-3.

Van Bussel, L.G.J., Ewert, F., Leffelaar, P.a., 2011. Effects of data aggregation on simulations of crop phenology. Agric. Ecosyst. Environ. 142, 75–84. http://dx.doi.org/10.1016/j.agee.2010.03.019.

Vereecken, H., Pachepsky, Y., Simmer, C., Rihani, J., Kunoth, A., Korres, W., Graf, A.,

Hendricks Franssen, H.J., Thiele-Eiche, I., Shao, Y., 2016. On the role of patterns in understanding the functioning of soil-vegetation-atmosphere systems. J. Hydrol. 542, 63–86. http://dx.doi.org/10.1016/j.jhydrol.2016.08.053.

van Wart, J., Kersebaum, K.C., Peng, S., Milner, M., Cassman, K.G., 2013. Estimating crop yield potential at regional to national scales. F. Crop. Res. 143, 34–43. http://dx.doi.org/10.1016/j.fcr.2012.11.018.

Wassenaar, T., Lagacherie, P., Legros, J., Rounsevell, M., 1999. Modelling wheat yield responses to soil and climate variability at the regional scale. Clim. Res. 11, 209–220. http://dx.doi.org/10.3354/cr011209.

de Wit, A.J.W., Boogaard, H.L., van Diepen, C.A., 2005. Spatial resolution of precipitation and radiation: the effect on regional crop yield forecasts. Agric. For. Meteorol. 135, 156–168. http://dx.doi.org/10.1016/j.agrformet.2005.11.012.

Zhang, Y., Feng, L., Wang, E., Wang, J., Li, B., 2012. Evaluation of the APSIM-wheat model in terms of different cultivars, management regimes and environmental conditions. Can. J. Plant Sci. 92, 937–949. http://dx.doi.org/10.4141/cjps2011-266.

Zhang, L., Yu, D., Shi, X., Xu, S., Xing, S., et al., 2014. Effects of soil data and simulation unit resolution on quantifying changes of soil organic carbon at regional scale with a biogeochemical process model. PLoS One 9 (2), e88622. http://dx.doi.org/10.1371/journal.pone.0088622.

Zhang, L., Zhuang, Q., Li, X., Zhao, Q., Yu, D., Liu, Y., Shi, X., Xing, S., Wang, G., 2016a. Carbon sequestration in the uplands of eastern China: an analysis with high-resolution model simulations. Soil Tillage Res. 158, 165–176. http://dx.doi.org/10.1016/j.still.2016.01.001.

Zhang, X., Xu, M., Sun, N., Xiong, W., Huang, S., Wu, L., 2016b. Modelling and predicting crop yield, soil carbon and nitrogen stocks under climate change scenarios with fertiliser management in the North China plain. Geoderma 265, 176–186. http://dx.doi.org/10.1016/j.geoderma.2015.11.027.

Zhao, G., Hoffmann, H., van Bussel, L., Enders, A., Specka, X., Sosa, C., Yeluripati, J., Tao, F., Constantin, J., Raynal, H., Teixeira, E., Grosz, B., Doro, L., Zhao, Z., Nendel, C., Kiese, R., Eckersten, H., Haas, E., Vanuytrecht, E., Wang, E., Kuhnert, M., Trombi, G., Moriondo, M., Bindi, M., Lewan, E., Bach, M., Kersebaum, K., Rötter, R., Roggero, P., Wallach, D., Cammarano, D., Asseng, S., Krauss, G., Siebert, S., Gaiser, T., Ewert, F., 2015a. Effect of weather data aggregation on regional crop simulation for different crops, production conditions, and response variables. Clim. Res. 65, 141–157. http://dx.doi.org/10.3354/cr01301.

Zhao, G., Siebert, S., Enders, A., Rezaei, E.E., Yan, C., Ewert, F., 2015b. Demand for multi-scale weather data for regional crop modeling. Agric. For. Meteorol. 200, 156–171. http://dx.doi.org/10.1016/j.agrformet.2014.09.026.

Zhao, G., Hoffmann, H., Yeluripati, J., Xenia, S., Nendel, C., Coucheney, E., Kuhnert, M., Tao, F., Constantin, J., Raynal, H., Teixeira, E., Grosz, B., Doro, L., Kiese, R., Eckersten, H., Haas, E., Cammarano, D., Kassie, B., Moriondo, M., Trombi, G., Bindi, M., Biernath, C., Heinlein, F., Klein, C., Priesack, E., Lewan, E., Kersebaum, K.-C., Rötter, R., Roggero, P.P., Wallach, D., Asseng, S., Siebert, S., Gaiser, T., Ewert, F., 2016. Evaluating the precision of eight spatial sampling schemes in estimating regional means of simulated yield for two crops. Environ. Model. Softw. 80, 100–112. http://dx.doi.org/10.1016/j.envsoft.2016.02.022.