

RESEARCH ARTICLE

Impact of Spatial Soil and Climate Input Data Aggregation on Regional Yield Simulations

Holger Hoffmann^{1*}, Gang Zhao¹, Senthold Asseng², Marco Bindi³, Christian Biernath⁴, Julie Constantin⁵, Elsa Coucheney⁶, Rene Dechow⁷, Luca Doro⁸, Henrik Eckersten⁹, Thomas Gaiser¹, Balázs Grosz⁷, Florian Heinlein⁴, Belay T. Kassie², Kurt-Christian Kersebaum¹⁰, Christian Klein⁴, Matthias Kuhnert¹¹, Elisabet Lewan⁶, Marco Moriondo¹², Claas Nendel¹⁰, Eckart Priesack⁴, Helene Raynal⁵, Pier P. Roggero⁸, Reimund P. Rötter¹³, Stefan Siebert¹, Xenia Specka¹⁰, Fulu Tao¹³, Edmar Teixeira¹⁴, Giacomo Trombi³, Daniel Wallach⁵, Lutz Weihermüller¹⁵, Jagadeesh Yeluripati¹⁶, Frank Ewert¹



OPEN ACCESS

Citation: Hoffmann H, Zhao G, Asseng S, Bindi M, Biernath C, Constantin J, et al. (2016) Impact of Spatial Soil and Climate Input Data Aggregation on Regional Yield Simulations. PLoS ONE 11(4): e0151782. doi:10.1371/journal.pone.0151782

Editor: David A Lightfoot, College of Agricultural Sciences, UNITED STATES

Received: December 24, 2015

Accepted: March 3, 2016

Published: April 7, 2016

Copyright: © 2016 Hoffmann et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All required input data for the manuscript is publicly available as under the following doi / link: <http://dx.doi.org/10.7910/DVN/CQJ5BB>.

Funding: This work was financially supported by the German Federal Ministry of Food and Agriculture (BMEL) through the Federal Office for Agriculture and Food (BLE), (2851ERA01J). FT and RPR were supported by FACCE MACSUR (3200009600) through the Finnish Ministry of Agriculture and Forestry (MMM). EC, HE, and EL were supported by The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (220-

1 Crop Science Group, INRES, University of Bonn, Bonn, Germany, **2** Agricultural & Biological Engineering Department, University of Florida, Gainesville, Florida, United States of America, **3** Department of Agri-food Production and Environmental Sciences, University of Florence, Florence, Italy, **4** Institute of Biochemical Plant Pathology, German Research Center for Environmental Health, Helmholtz Zentrum München, Neuherberg, Germany, **5** INRA, Auzeville, France, **6** Department of Soil and Environment, Swedish University of Agricultural Sciences, Uppsala, Sweden, **7** Thünen-Institute of Climate-Smart-Agriculture, Braunschweig, Germany, **8** Desertification Research Group, Università degli Studi di Sassari, Sassari, Italy, **9** Department of Crop Production Ecology, Swedish University of Agricultural Sciences, Uppsala, Sweden, **10** Institute of Landscape Systems Analysis, Leibniz Centre for Agricultural Landscape Research, Müncheberg, Germany, **11** Institute of Biological and Environmental Sciences, School of Biological Sciences, University of Aberdeen, Aberdeen, Scotland, United Kingdom, **12** CNR-Ibimet, Florence, Italy, **13** Environmental Impacts Group, Natural Resources Institute Finland (Luke), Vantaa, Finland, **14** Systems Modelling Team (Sustainable Production Group), The New Zealand Institute for Plant and Food Research Limited, Canterbury Agriculture & Science Centre, Lincoln, New Zealand, **15** Agrosphere Institute (IBG-3), Forschungszentrum Jülich GmbH, Jülich, Germany, **16** The James Hutton Institute, Craigiebuckler, Aberdeen, Scotland, United Kingdom

* hhoffmann@uni-bonn.de

Abstract

We show the error in water-limited yields simulated by crop models which is associated with spatially aggregated soil and climate input data. Crop simulations at large scales (regional, national, continental) frequently use input data of low resolution. Therefore, climate and soil data are often generated via averaging and sampling by area majority. This may bias simulated yields at large scales, varying largely across models. Thus, we evaluated the error associated with spatially aggregated soil and climate data for 14 crop models. Yields of winter wheat and silage maize were simulated under water-limited production conditions. We calculated this error from crop yields simulated at spatial resolutions from 1 to 100 km for the state of North Rhine-Westphalia, Germany. Most models showed yields biased by <15% when aggregating only soil data. The relative mean absolute error (rMAE) of most models using aggregated soil data was in the range or larger than the inter-annual or inter-model variability in yields. This error increased further when both climate and soil data were aggregated. Distinct error patterns indicate that the rMAE may be estimated from few soil

2007-1218) and by the strategic funding 'Soil-Water-Landscape' from the faculty of Natural Resources and Agricultural Sciences (Swedish University of Agricultural Sciences) and thank professor P-E Jansson (Royal Institute of Technology, Stockholm) for support. JC, HR, and DW thank the INRA ACCAF metaprogram for funding and Eric Casellas from UR MIAT INRA for support. CB was funded by the Helmholtz project "REKLIM - Regional Climate Change". CK was funded by the HGF Alliance "Remote Sensing and Earth System Dynamics" (EDA). FH was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) under the Grant FOR1695. FE and SS acknowledge support by the German Science Foundation (project EW 119/5-1). HH, GZ, SS, TG, and FE thank Andreas Enders and Gunther Krauss (INRES, University of Bonn) for support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

variables. Illustrating the range of these aggregation effects across models, this study is a first step towards an ex-ante assessment of aggregation errors in large-scale simulations.

Introduction

Process-based crop models are increasingly used for large-scale simulations at regional [1] and global scale [2]. However, these models have typically been developed at the field scale, where model driving variables are easily obtained [3–4]. Using crop models at scales other than those they were developed for may lead to biased simulations for two reasons. Firstly, biophysical processes or phenomena often depend on the scale [5–6]. Models built to simulate these processes at a given scale may therefore not be valid when changing the scale [7]. Secondly, changing the spatial resolution by aggregating or disaggregating data bears the risk of biased simulations due to modified data [8–10]. This is the so-called nonlinear aggregation error [8] or aggregation effect [11–12]. In practice, simulation studies often rely on data of different types and resolutions, disaggregated or aggregated by different methods to one resolution [13–14]. The error associated with this practice is, however, rarely considered when assessing model validity [15–16].

Large-scale yield simulations with process-based crop models rely heavily on climate and soil data inputs. Spatial resolution and combination of these inputs are known to have a large impact on the model output [17]. However, these inputs are often only available at low spatial resolution or are aggregated by spatially averaging climate [18–19] and selecting the soil type with the largest spatial coverage within a larger area [20–21]. Despite its relevance, few studies have characterized the implications of this practice (Table 1). Most of these studies have focused on climate data aggregation. For crop yield, they report an aggregation error of up to 4.7% root mean square deviation (RMSD). However, [11–12] showed that this error differs across models. Additionally, the differences mainly depend on the analyzed spatial and temporal resolution and aggregation method. While this hampers a fair comparison across studies, the aggregation error may further differ between response variables of the same simulation study. Errors considerably larger than the reported regional biases (Table 1) may therefore occur under different conditions. For instance, [10] reported errors in annual net primary production (NPP) of up to 18%, using an environmental model. As a consequence, data aggregation poses an additional complex problem when estimating model outcome uncertainty.

Unlike climate data aggregation, the error in crop yield in combination with soil input data aggregation has only been investigated by few studies. This is in contrast to the fact that soils can be a major source of spatial yield variability in temperate climate [22] and thus a source of potentially large aggregation effects. Additionally, interactions of climate and soil data aggregation have rarely been studied. Using different approaches, [23–25] analyzed the impact of aggregating both soil, climate, as well as partly their interactions for single models (EPIC, CLIMCROP, and LINTUL2). They concluded that 10 km (USA, Denmark) and 100 km (Germany) resolution are sufficient to reproduce regional yield statistics. However, the relative importance of soil and climate data for the regional yield bias differed among these studies. [23] state that soil resolution had no effect on model performance in the Great Plains (USA), while [24] found decreasing winter wheat yields and increasing irrigation demand in Denmark with decreasing soil resolution. Following input data aggregation, [25] report differences in drought stress at resolutions of up to 100 km as compared to 1 km resolution. Drought stress was overestimated in drought prone regions and underestimated in more humid regions.

Table 1. Publications focusing on the effect of spatial input data aggregation on crop and environmental model output variables. Input data variables: climate (c), soil (s), phenology (p), management (m), topography (t), land-use (lu), vegetation (v). Aggregation methods: spatial average (av), area majority (m), direct use of maps at given resolution (map), other/various (v). Crops: winter wheat (ww), silage maize (sm), grain maize (gm), spring barley (sb). Model type: crop (c), ecosystem (e), energy balance (r).

Region ^a	Aggregation		Resolution or spatial scale	Model,no.	Response variable	Error	Reference
	data	method					
FI	c	av	10..100 km	c, 4	sb yield	Bias ^{c,d,e} ≤1.5%	[15]
DE, FR	c, s	av, map	10..50 km	c, 1	gm, ww yield	RMSE ^d ≤4.7%	[26]
AU	c ^g	av	0.12..0.96 km	r, 1	ETa ^h	Bias ^c ≤15%	[27]
NRW ^b	c	av	1..100 km	c, 13	sm, ww yield	Bias ^c ≤ 3.5%, RMSE ≤4.5%	[11]
NRW ^b	c	av	1..100 km	c, 13	sm, ww yield	Bias ^c ≤ 3.5%, RMSE ≤4.5%	[12]
DE	c	av	1..100 km	c, 1	ww yield	RMSE ^d ≤5%	[28]
NRW ^b	s	map	1:5·10 ⁴ ..1:10 ⁶	c, 4	ww yield	Bias ^{c,d,e} ≤2.9%	[16]
DE	s	map	1:5·10 ³ ..1:10 ⁵	c, 1	nitrogen leach.	Bias ^{c,d} ≤8.6%	[29]
US	c, s	v	0.5°..2.8°	c, 1	ww, gm yield	MAE ≤0.29 (ww), ≤0.7 t ha ⁻¹ (gm)	[23]
DK	c, s	v	various	c, 1	ww yield	Bias ^{c,d} ≤14%	[24]
DE	c, s, p	av, m	1..100 km	c, 1	ww heat stress	Bias = 0.003 ^b [-]	[25]
USA	c, s, lu	av, m, v	7.5..45 km	c, 1	gm yield	Bias ^{c,d} ≤2.3%	[1]
US	c, s, t, v	av	1..110 km	e, 1	NPP ^f	Bias = 18%	[10]
DE	c, m	av	10..100 km	c, 1	ww phenology	RMSE ≤4.4%	[30]

^a) country code (ISO 3166–1 alpha-2).

^b) North Rhine-Westphalia, DE.

^c) average over years, region and models.

^d) calculated from reference.

^e) based on median yield.

^f) net primary production.

^g) from remote sensing.

^h) actual evapotranspiration.

doi:10.1371/journal.pone.0151782.t001

These findings highlight the importance of interactions between model and input data on aggregation effects. Consequently, changing model or region may result in different conclusions. [10] estimated the contribution of different data types to the regional bias of NPP due to aggregating by spatial averaging from 1 to 110 km resolution. This bias was 32% due to topography, 32 to 50% due to climate, and 17 to 34% due to averaging spatial variations in vegetation and soil water holding capacity (SWHC). However, as a direct input to the employed environmental model, [10] averaged SWHC. In contrast, most crop models rely on layer-defined soil profiles (e.g. texture and physical properties) and spatial averaging is not a suitable method. Furthermore, [10] calculated the contribution of the different data types to the regional bias by averaging each data type to a single spatial average. Unfortunately, while this approach does not reflect the current practices in large-scale crop simulations, it also does not allow showing changes in spatial patterns due to data aggregation. A comprehensive analysis of soil and climate input data aggregation is therefore essential to establish a base for evaluating the significance of crop simulation studies which involve changes in scale. [23] pointed out, that such analysis should be extended to more than one model. Nevertheless, a systematic multi-model comparison of the effects of soil and climate data aggregation as well as the interaction between soil and climate data aggregation has not been carried out so far.

This work aims at analysing the effects of spatially aggregating crop model input data (climate and soil) on simulated yield for a range of crop models. We hypothesize that, under

average conditions, the aggregation effect increases with increasing drought and decreasing plant available water capacity (awc) of a soil, as suggested by [7, 16]. Finally, we assume that cropping regions characterized by the occurrence of extreme soil (very low / high water holding capacity) and climate conditions (very dry / moist) are prone to the largest aggregation errors.

Material and Methods

Procedure and regional focus

We tested the hypothesis given above in the federal state of North-Rhine Westphalia (NRW), Germany (Fig 1A). NRW has an area of 34,098 km² [31] with elevations <1000 m above sea level (Fig 1B) and a humid, temperate climate. Prevailing soil types are Cambisols, Luvisols and Stagnosols (FAO key reference soil groups: CM, LV, ST, respectively; [32]). The heterogeneous topography results in several agro-ecological zones with different soils and temperature-precipitation regimes. Spatial aggregation effects were assessed by aggregating soil and climate data to spatial resolutions varying between 1 km and 100 km. The obtained data was then used as model input for simulations of winter wheat (*Triticum aestivum* L.) and silage maize (*Zea mays* L.). Aggregation effects were estimated by relating simulated yields to climate and soil input.

Climate and soil data processing / aggregation methods

For climate data, time series of minimum, mean and maximum air temperature (2 m above ground), precipitation, global radiation, wind speed and relative humidity were used. For this purpose, daily records from 280 weather stations and an interpolated grid of 1 km resolution of monthly records were obtained from the German Meteorological Service for the period 1982 to 2011. The station records of daily temporal resolution and the monthly grids of 1 km spatial resolution were combined to daily time series of 1 km spatial resolution following [28]. The climate properties of the region for the different spatial resolutions were given by [11]. In addition, Fig 2 shows the inter-annual variability of the climatic water balance (cwb) calculated as precipitation (P) minus reference evapotranspiration (ET): $cwb = P - ET$. ET was calculated according to [33].

For soil data, combinations of soil types at the scale of 1:50,000 [34] and physical parameters were joined to mapping units [16, 35]. Further soil parameters were obtained as follows: 1) Topsoil organic carbon and pH were taken from the database FIS StoBo [36]; 2) Organic carbon and C:N-ratio of deeper soil layer was approximated using pedotransfer functions [16, 35]; 3) Top soil layer C:N-ratio was set to 10; 4) Volumetric gravel content and gravel content corrected bulk density were approximated following [37–38]; 5) Soil surface albedo was estimated from soil organic carbon (Equation A in S1 File).

Climate and soil data were spatially aggregated in order to obtain grids of 1, 10, 25, 50 and 100 km resolution. For this purpose climate data of spatial resolution of 1 km were spatially averaged. Soil data was aggregated by selecting the dominant soil at approximately 300 m resolution raster within each grid cell of coarser resolution. A dominant soil was chosen by selecting the mapping unit of the largest area coverage (area majority). The main soil types and properties across resolutions are shown in Fig 3 and Table 2. A complete list of aggregated model input data is given by Table A in S1 Table.

In addition to the described grids with a combination of soil and climate input at the respective resolution, simulations were carried out using the averaged climate time series of the entire region (mean regional climate) as well as using only one soil type representative of a typical crop land of the region (one representative soil).

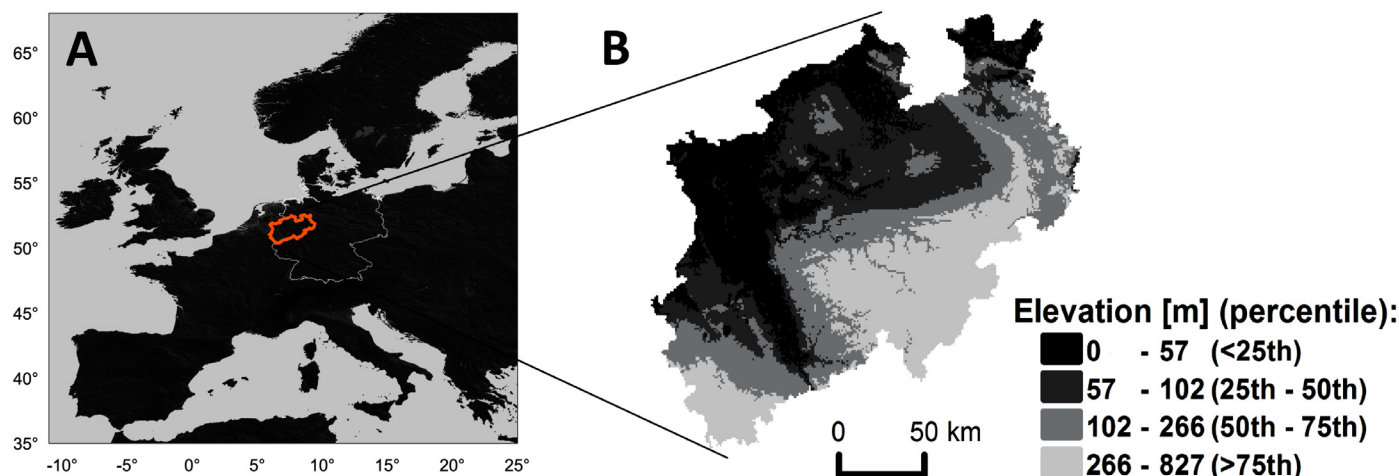


Fig 1. The state of North-Rhine Westphalia (NRW). (A) Location in Europe (orange line: border of NRW; white line: border of Germany). (B) Elevation above sea level.

doi:10.1371/journal.pone.0151782.g001

All described crop model input data is available under the CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE [39].

Crop simulations

Ensembles of models were used to simulate winter wheat (11 models) and silage maize (9 models) crop growth. These models are currently used in addressing different research questions at various scales and are described in Table 3. The models simulated the phenology, growth and yield of the crops for the period 1982 to 2011. Here, yield of winter wheat refers to grain yield

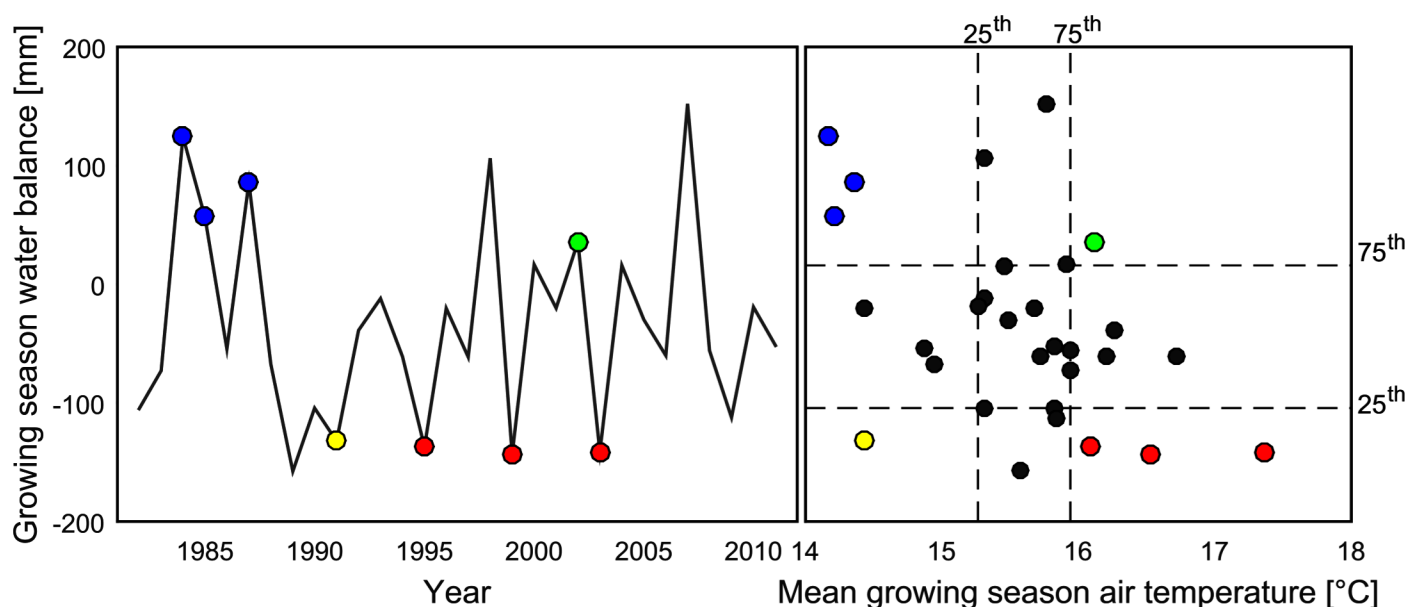


Fig 2. Inter-annual variability of growing season temperature (T_{av}) and climatic water balance of the region of North Rhine-Westphalia (regional median, calculated from 1 km resolution). Extreme years of the time series are highlighted: cold and dry (yellow), cold and wet (blue), hot and dry (red) and warm and wet (green). Extreme years are years below the 25th (cold, dry) or above the 75th (hot, wet) percentile of mean annual temperatures and annual precipitation, respectively.

doi:10.1371/journal.pone.0151782.g002

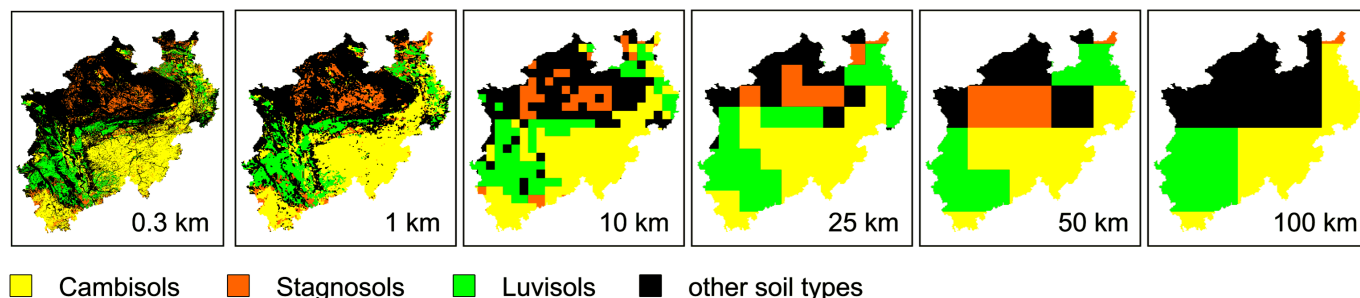


Fig 3. Spatial distribution of main soil types in North Rhine-Westphalia as influenced by aggregation. Resolutions of 1 to 100 km were aggregated from the source data of 0.3 km spatial resolution.

doi:10.1371/journal.pone.0151782.g003

whereas silage maize yield is aboveground biomass. Furthermore, both crops were evaluated for water-limited production situations, limited by crop characteristics, atmospheric CO₂, temperature, radiation, precipitation and other soil water balance components [40–41]. A constant management except of harvest dates was used for all grid cells (Table B in S1 Table). Models were calibrated at 1 km resolution, using one typical sowing and one typical harvest date per crop in addition to the whole region area weighted average of observed yields. The latter was derived from 1999 to 2011 and from 2000 to 2008 county statistics for winter wheat and silage maize, respectively [42].

Simulations were carried out for each grid cell for each of the following five combinations of soil and climate data grids: I) soil resolutions 1 to 100 km x mean climate; II) one select soil x climate resolutions 1 to 100 km; III) soil resolutions 1 to 100 km x climate at 1 km resolution; IV) soil at 1 km resolution x climate resolutions 1 to 100 km; V) soil x climate at same resolutions 1 x 1, 10 x 10, 25 x 25, 50 x 50 and 100 x 100 [km x km].

Analysis

Simulated yields at resolutions >1 km were spatially disaggregated to 1 km resolution prior to further analysis. Subsequently, we calculated the agreement of the models with respect to the average yields of each cell, using an agreement indicator (Equations E and F in S1 File). Model agreement in time (years) and space (grid cells) was portrayed via a Taylor diagram [71]. Probability density functions (pdf) for yield were obtained by kernel density estimation using a Gaussian kernel (see [72] for equations). The regional bias (\bar{Y}_d) and relative mean absolute error (rMAE) of yields at coarser resolutions to yields at 1 km resolution were computed according to [11], Equation D in S1 File and Equation 4 in [12], respectively. These average statistics were based on single cell—single year differences of disaggregated coarser resolutions to

Table 2. Main soil types in North Rhine-Westphalia as influenced by aggregation.

Resolution	No. of soil types	Cambisols	Stagnosols	Luvisols	Cambisols, Stagnosols and Luvisols
[km]	[–]	Fraction of total area [%]			
0.3	92	29.7	9.8	12.8	52.2
1	69	33.0	10.4	15.4	58.8
10	28	36.9	10.0	21.3	68.2
25	14	37.2	10.9	32.3	80.4
50	7	36.0	15.2	25.5	76.7
100	5	32.6	0.5	24.4	57.6

doi:10.1371/journal.pone.0151782.t002

Table 3. Models.

No.	Model	References
1	AgroC	[43]
2	APSIM-NWHEAT ^a	[44–46]
3	APSIM	[46–47]
4	Century	[48–49]
5	COUP ^a	[50–52]
6	CROPSYST ^b	[53]
7	DailyDayCent	[54–57]
8	ExpertN-SPASS ^{a,b}	[58–59]
9	EPIC v. 0810	[60]
10	HERMES	[61–62]
11	MCWLA ^{a,b}	[63–64]
12	MONICA	[65–66]
13	SIMPLACE<LINTUL5;SLIM>	[3, 67]
14	STICS	[68–70]

^a only simulated wheat.

^b only simulated yield at 1 km resolution.

doi:10.1371/journal.pone.0151782.t003

a 1 km resolution. In order to obtain better insights on the direction and distribution of errors we additionally analysed these single cell—single year differences as follows:

$$\Delta_j = x'_j - x_j \quad (1)$$

$$rAE_j = |\Delta_j| \cdot \bar{x}^{-1} \quad (2)$$

where Δ_j is the difference of an aggregated variable of interest disaggregated to 1 km resolution (x'_j) to the variable of interest at 1 km resolution (x) at a grid cell (j). rAE_j is the relative absolute error, relative to the regional average of the variable of interest at 1 km resolution (\bar{x}).

In order to test how soil properties and climate regimes interact with aggregation effects, we first analysed the soil and climate input variables which explained the yield variance best over the region. Subsequently, further analysis was subdivided based on these variables. Best explaining variables were analysed by regression of average yields (1982 to 2011, median of models) of the 34168 grid cells at 1 km resolution to all possible combinations of four soil and climate input variables via PLS-regression [73]. Based on these findings (see [result](#) section), further analysis was refined for plant available water capacity awc, mean growing season temperature Tav and growing season climatic water balance cwb ([Fig 2](#)). Therefore, yields and aggregation effects were analysed separately for low and high values of the cwb and awc or Tav. High and low values were extracted by selecting values below the 25th or above the 75th percentile. The cwb was computed according to [33] from 1 km resolution climate data and validated by visual comparison with published maps [74] and comparison with regional mean time series of annual values ($R^2 > 0.99$, Bias = −8.4 mm) with [75].

Ethics Statement

The study did not require specific permissions. All research was carried out with publicly available data. We confirm that the field studies did not involve endangered or protected species.

Results

Soil and climate input data

Spatial aggregation decreased the spatial variance across climate time series (Table C in [S1 Table](#)). Averaging out climate extremes narrowed regional extremes, e.g. decreasing regional maximum annual precipitation or increasing minimum air temperature. In contrast, soil aggregation by means of majority increased the spatial variance of soil variables across the grid cells of the region (Table C in [S1 Table](#)). Despite a constantly decreasing number of soil types with aggregation ([Table 2](#)), soil variables regional minimum and maximum changed abruptly with changing spatial composition of soil types. For both aggregation methods, area mean and median were less affected (Table C in [S1 Table](#)).

Simulated crop yields

Simulated crop yields varied across crops, years, and climate and soil conditions and in the resulting spatial pattern, as well as across models. Following we show this variability for each of these dimensions at 1 km resolution.

Crop. Regional crop yields (regional median, average over years) from 1 km resolution of both climate and soil data were 7.2 and 15.2 t ha⁻¹ for winter wheat and silage maize, respectively ([Table 4](#)). Instead, using either the regional climate data averages or the single select soil as model input led to yields, up to 0.9 and 0.8 t ha⁻¹ larger in the regional median for wheat and maize, respectively.

Temporal variability. The regional yields showed an inter-annual variability (mean of year-to-year standard deviation from each grid cell and model) of 1.1 t ha⁻¹ (15.3%) and 2.4 t ha⁻¹ (15.8%) for winter wheat and silage maize, respectively. Yields from simulations with spatially averaged climate (1.0 and 2.2 t ha⁻¹) or with one select soil (0.8 and 1.9 t ha⁻¹) showed an inter-annual yield variability of a similar range. In the case of winter wheat, however, regional yields ranged mainly from 7 to 8 t ha⁻¹ with considerably lower yields in 2010 and 2011 (Figure B in [S1 Fig](#)).

Spatial variability. We found, that the spatial yield variability could be explained by 68% (winter wheat) and 81% (silage maize) based on four variables (Figure A in [S1 Fig](#)): growing season precipitation, plant available water capacity of the soil profile (awc, [mm]), soil profile depth and topsoil awc in the case of wheat and growing season mean daily temperature, awc, soil profile depth and topsoil awc in the case of silage maize. The single variable explaining most of the spatial variability was awc, (38% wheat, 58% maize), followed by soil profile depth (33% wheat, 36% maize). Climate variables showed low explanation power for spatial variability for wheat (<10%), whereas spatial variability of average silage maize yields were explained 31, 35 and 36% by average daily minimum, mean and maximum air temperature, respectively. As a consequence, the further analysis was refined (see [methods](#)) for awc, Tav and cwb of the growing season.

In contrast to the year-to-year fluctuation in yield, the spatial variability in yields depended more on the spatial input data resolution. For instance, the mean of cell-to-cell standard deviation from each year and model from both climate and soil input data at 1 km resolution was 1.3 and 3.1 t ha⁻¹, for winter wheat and silage maize, respectively. These values decreased when using the regional climate data averages (1.0 and 2.2 t ha⁻¹) or the single select soil (0.6 and 1.7 t ha⁻¹) as model input. Thus, substituting high resolution input data by a single soil decreased the spatial variability more than substitution by a single climate time series.

Single year crop yields ranged from 2.5 to 9.5 t ha⁻¹ and from 4 to 20 t ha⁻¹ for winter wheat and silage maize respectively, when averaged across models (Figure C in [S1 Fig](#)). While silage

Table 4. Simulated water-limited yield at 1 km resolution over the period 1983 to 2011 and corresponding bias (\bar{Y}_d) of coarser resolutions to the yield at 1 km resolution. Values were calculated for each model first. Thereafter minimum (min), median and maximum (max) values across models were calculated. ww: Winter wheat; sm: Silage maize.

Crop	Yield [t ha ⁻¹]					\bar{Y}_d [%]											
	Model input*		1 km			10 km			25 km			50 km			100 km		
	Agr.	Constant	min	median	max	min	median	max	min	median	max	min	median	max	min	median	max
ww	s	c1	6.1	7.2	7.7	-0.6	0.3	1.8	-2.9	-1.2	0.8	0.1	0.8	6.3	-8.8	-0.1	17.3
ww	c	s1	6.1	7.2	7.7	-0.8	-0.2	1.6	-0.1	0.3	2	0.0	0.9	2.3	0.3	1.7	3.8
ww	s, c	-	6.1	7.2	7.7	-1.2	0.1	1.9	-2.6	-0.6	1.9	0.9	1.9	7.6	-7.9	1.3	17.8
ww	s	cNRW	6.1	7.6	8.4	-0.7	0.0	2.8	-3.3	-1.1	0.7	-1.8	0.9	3.7	-5.0	-0.4	16.9
ww	c	sNRW	7.0	8.1	9.3	-0.5	0.0	0.6	0.1	0.7	1.2	0.5	1.1	2.1	0.7	1.1	3.3
sm	s	c1	10.4	15.2	17.7	-0.8	0.2	1.6	-2.3	-0.7	1.3	0.1	1.2	5.2	-7.4	0.9	42.7
sm	c	s1	10.4	15.2	17.7	-1.0	-0.1	3.3	-0.7	0.7	3.8	-0.2	1.4	4.5	-3.1	2.3	6.1
sm	s, c	-	10.4	15.2	17.7	-1.3	0.4	3.5	-2.1	0.7	3.1	0.9	3.1	7.9	-5.7	2.1	18.5
sm	s	cNRW	11.3	15.3	19.8	-0.8	-0.1	2.1	-2.5	-1.0	0.6	-0.6	0.7	4.3	-2.6	0.9	16.7
sm	c	sNRW	13.1	16.0	20.5	-1.0	0.1	0.8	-0.6	0.3	1.2	-0.3	0.6	1.7	-1.5	0.9	2.6

*) Abbreviations refer to soil (s, s1, sNRW) or climate (c, c1, cNRW) model input. Agr: Aggregated to 10, 25, 50 or 100 km resolution. Constant: Not aggregated model input kept constant at 1 km resolution (c1, s1) or using an average climate time series (cNRW) or one soil for the region (sNRW). Example: Values for aggregated soil (s) and constant climate at 1 km resolution (c1) are min/median/max yields from simulations with soil resolutions of 10 to 100 km in combination with climate data of 1 km resolution.

doi:10.1371/journal.pone.0151782.t004

maize yields for years with a growing season cwb above -200 mm increased with increasing awc, winter wheat yields continuously increased with both awc and cwb. In this context, winter wheat yields above 8.5 t ha⁻¹ were mainly reached above 300 mm growing season cwb and on soils with awc >100 mm. Consistently, winter wheat yields below 7.5 t ha⁻¹ occurred mainly with negative cwb and on soils with awc <200 mm. Silage maize yields of 16 t ha⁻¹ were achieved on soils with awc >200 mm, regardless of cwb. Accordingly, yields showed characteristic probability density functions at upper and lower quartiles of awc and cwb (Fig 4).

The interaction of soil and climate conditions was reflected in the spatial pattern of the crop yields (Fig 5). With 8.7% and 90.6% of all grid cells at 1 km resolution exhibiting soil profile depths of ≤0.6 m and ≥1.7 m, respectively, the comparatively small fraction of shallow soils (≤0.6 m) led to characteristic spatial yield patterns. For instance, areas of low yield were found in the north-east, central south, and south-west of NRW, coinciding with shallow soils. These shallow soils exhibited low awc and were mainly located at higher elevations (Fig 1). Unlike these differences in specific soil variables of a given soil type, average yields across soil types differed less. From the three main soil types of the region (Fig 3), the average yields over years were (mean ± standard deviation over the region) 7.1 ± 0.8, 7.5 ± 0.5, 7.2 ± 0.7 t ha⁻¹ for winter wheat and 13.9 ± 1.8, 15.9 ± 1.1, 14.7 ± 1.3 t ha⁻¹ for silage maize, respectively for Cambisols, Luvisols and Stagnosols. Highest average yields were found on gleyed Cambisols and Chernozem-Luvisols (>8 t ha⁻¹ winter wheat; >16.9 t ha⁻¹ silage maize). In contrast, yields below 6 t ha⁻¹ (winter wheat) and 12 t ha⁻¹ (silage maize) were found on a variety of soil types, mainly characterized by shallow profile depth and/or large gravel content.

Model variability. In general, models showed a larger coefficient of variation (cv) in silage maize yields (cv = 27%) than for winter wheat yields (cv = 18%). This is illustrated by the scatter of models in Taylor plots (Figure D in S1 Fig). Average yields of years and region ranged from 6.1 to 7.7 t ha⁻¹ and from 10.4 to 17.7 t ha⁻¹ for winter wheat and silage maize,

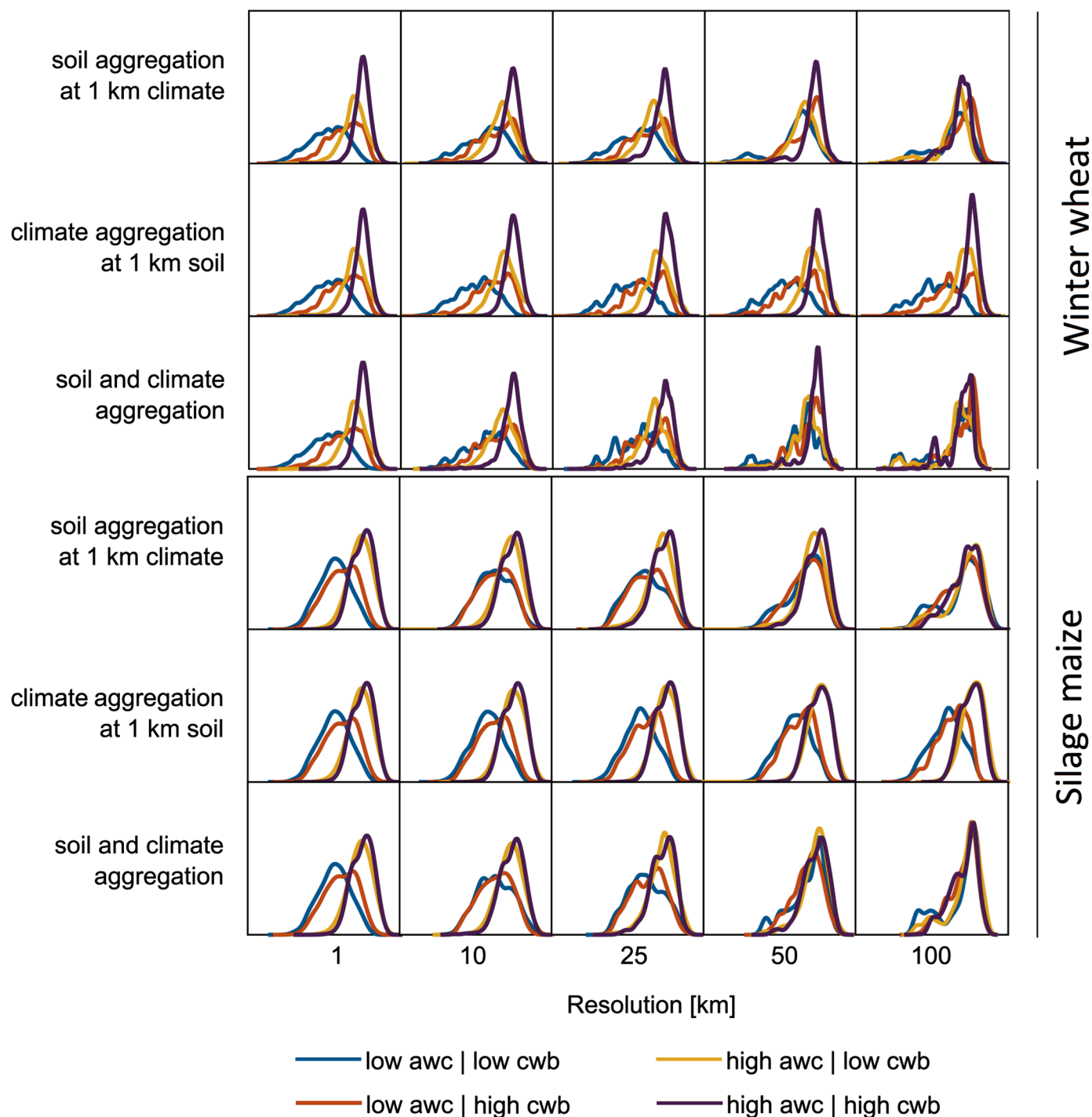


Fig 4. Simulated probability density function (pdf) of winter wheat and silage maize yield under water-limited conditions in North Rhine-Westphalia, Germany (1983–2011) as related to the aggregation of different data types as well as the plant available water capacity (awc) over the profile and climatic water balance (cwb) of the growing season. 'low' and 'high' refer to values below the 25th or above the 75th percentile. Pdf was estimated via kernel density (Gaussian kernel with bandwidth of 0.1 and 0.5 for wheat and maize, respectively). Subplots range from 2 to 10 t ha⁻¹ and 1 to 21 ha⁻¹ for wheat and maize, respectively.

doi:10.1371/journal.pone.0151782.g004

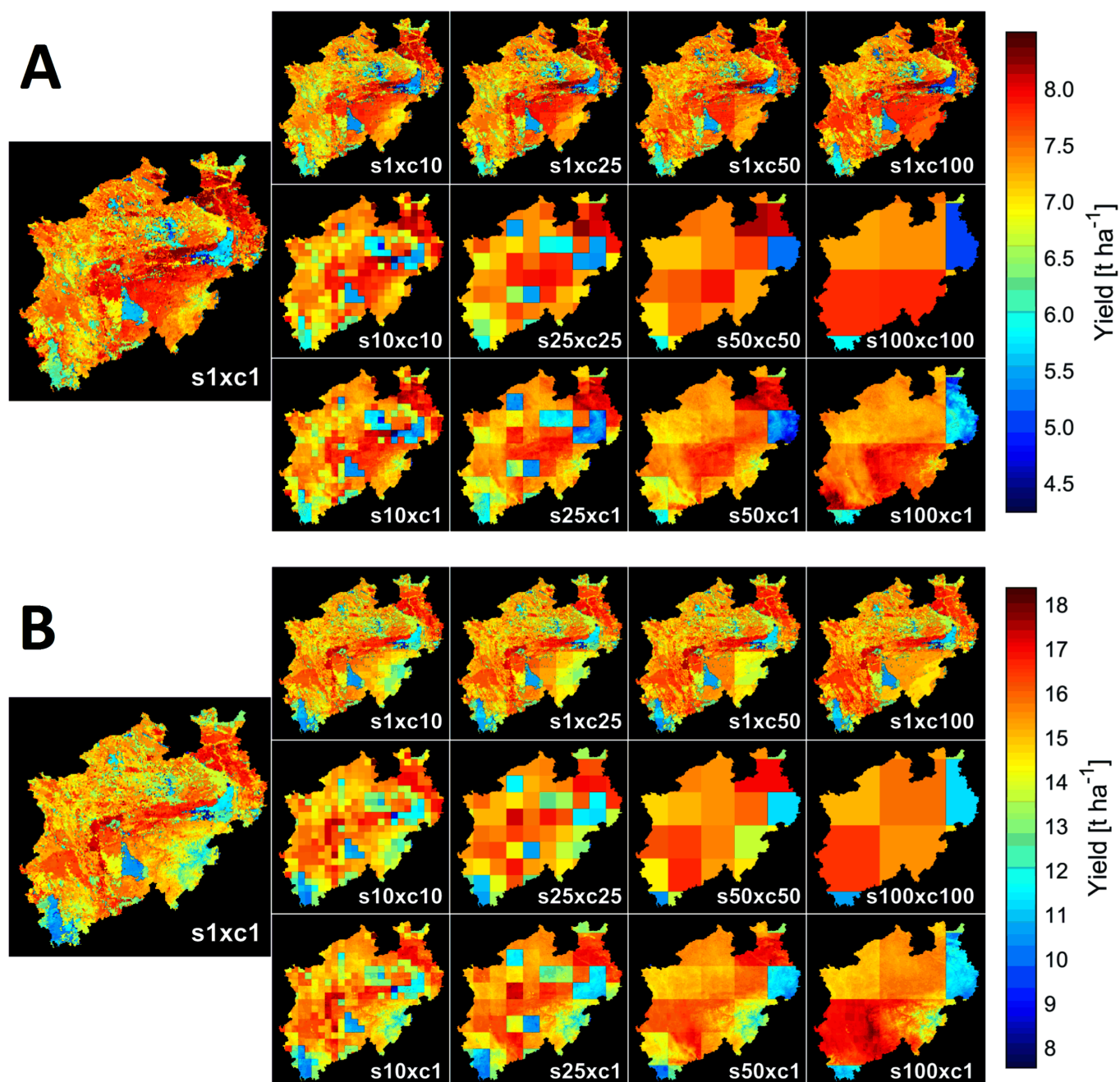


Fig 5. Simulated water-limited yield of North Rhine-Westphalia, Germany (1983–2011; mean of years and models). (A) Winter wheat. (B) Silage maize. Results are shown for a range of model input data combinations of soil (s) and climate (c) data of 1, 10, 25, 50 and 100 km resolution.

doi:10.1371/journal.pone.0151782.g005

respectively. In addition, the variability across models for single years and cells was generally comparable or larger than the inter-annual variability or variability of yields across the region. For instance, the standard deviation across crop yields for each cell and year was 1.4 and 4.1 t ha⁻¹ on average for winter wheat and silage maize, respectively. However, in contrast to the spatial variability, model variability did only decrease when using one single soil as model input (1.0 and 3.4 t ha⁻¹), but not when using an average climate time series as input (1.4 and 4.0 t

ha⁻¹). In general, models agreed better with increasing yield. This is partially reflected in the spatial pattern of model agreement (Figure E in [S1 Fig](#)). Characteristic patches of shallow soils can be distinguished as linked to low model agreement. Thus, models agreed less on winter wheat yields and silage maize biomass on shallow soils.

Yield aggregation effects

Yields simulated with aggregated input data differed from yields simulated at 1 km resolution ([Table 4](#)). The regional yield bias (\bar{Y}_A) was up to $\pm 3.1\%$ of yields at 1 km resolution (median of models, [Table 4](#)), but differed considerably under specific conditions as shown in the following.

Crop-specific aggregation effects. Relative to the regional mean yield at 1 km resolution, maximum \bar{Y}_A for both crops were comparable in the median of models. In more detail, differences between the two crops were found for single resolutions and input data combinations. For instance, \bar{Y}_A of opposite signs were found between the crops with both, soil and climate data, aggregated to 25 km resolution. However, differences in model median biases across all resolutions were relatively small as compared to the spread across models (see min and max \bar{Y}_A , [Table 4](#)).

Temporal variability. Regional biases varied less than 0.2 t ha⁻¹ ($\leq 1.6\%$; standard deviation) across years, regardless of resolution or data input combination (see also Figure B in [S1 Fig](#)). However, the inter-annual variability of single cells' yield deviation from yields at 1 km resolution was in the range of the single cells inter-annual variability at 1 km resolution. In this case, the standard deviation across all years was 0.8 and 1.7 t ha⁻¹ for winter wheat and silage maize, respectively, as compared to the standard deviation of yields of 1.1 and 2.4 t ha⁻¹, respectively. However, the rMAE was increased in years of lowest growing season cwb as compared to those of highest cwb in all input data combinations with aggregated soil data ([Fig 6](#), Figure F in [S1 Fig](#)). This larger rMAE in years with low growing season cwb was found independent of the growing season temperature.

Spatial variability. Similar to yields, also the rMAE varied considerably with type of input data. Three distinct effects were identified. Firstly, the probability density function (pdf) of yields from soils of low awc change when aggregating soil ([Fig 4](#)). At aggregations of ≥ 50 km, these pdf's resemble those from high awc soils. In contrast, aggregation of climate did not show this effect. Secondly, the aggregation error increased with decreasing awc when aggregating soil data ([Fig 7](#)). Further, when aggregating soil, cwb showed no clear impact on the aggregation error. However, using aggregated climate data in combination with 1 km soil data as model input showed a contrasting pattern. In this case, aggregation effects increased with increasing cwb and with decreasing awc at 1 km resolution. These two effects partially resulted in a third effect: Aggregation effects showed distinct functions depending on the input data type and resolution ([Fig 8](#)). Aggregated soil data combined with either climate data at 1 km or averaged over the region revealed the fraction of the rAE with nearly zero error (rAE $\sim 0\%$). For instance, the rAE at 10 km resolution was close to 0% in approximately $\geq 50\%$ of all cells and years. This fraction decreased with further aggregation. In contrast, climate data aggregation caused a larger fraction of low rAE ($0 \leq \text{rAE} \leq 10\%$). Consistently, soil aggregation led to a larger fraction of high rAE. Finally, the combined aggregation of soil and climate data led to distribution of rAE with both a larger fraction of low rAE as compared to soil data aggregation as well as a larger fraction of high rAE as compared to climate data aggregation.

Model variability. Crop models differed in their aggregation effects. While most models showed aggregation effects of similar range for a given condition, few models showed considerably larger errors under specific conditions. This is expressed in the minimum and maximum model biases ([Table 4](#)) as well as in the outliers when analysing the rMAE ([Fig 6](#)). However,

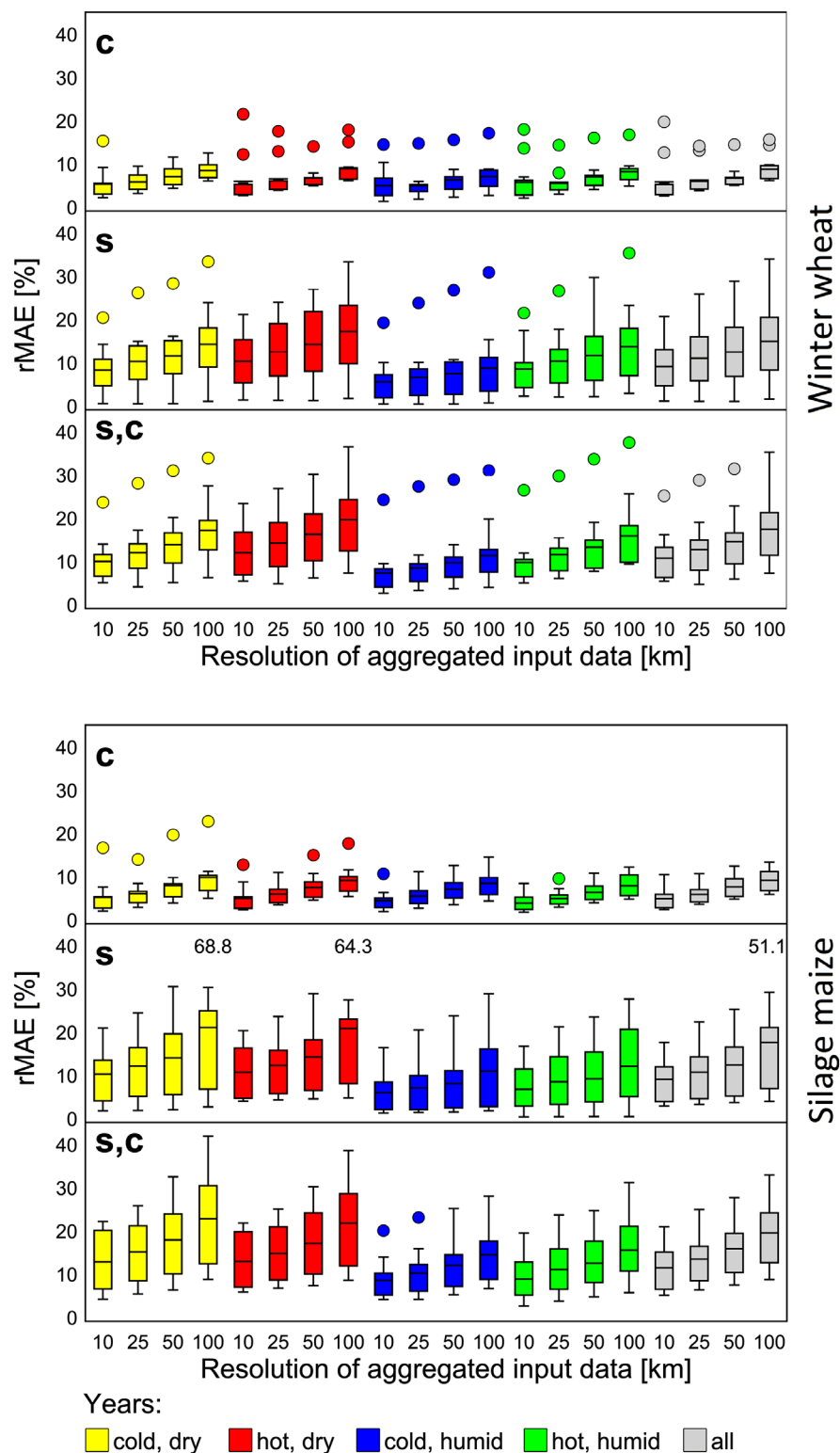


Fig 6. Relative mean absolute error (rMAE) of simulated winter wheat and silage maize yield under water-limited conditions for different spatial resolutions of model input data. c: aggregated climate x 1 km soil; s: aggregated soil x 1 km climate; s,c: aggregated soil x aggregated climate. The rMAE was calculated from data of extreme years (see Fig 2) and of all single years. Boxplots show the rMAE calculated from single model results (winter wheat: $n = 11$; silage maize: $n = 9$). The middle line indicates the mean

rMAE across models. Whiskers are Tukey style and extent to 1.5 times the interquartile range (see [76]). Extreme outliers are indicated by written numbers.

doi:10.1371/journal.pone.0151782.g006

these extreme results strongly depended on the specific condition (data type, resolution, year, crop, etc.).

Discussion

Data aggregation effects on regional soil and climate statistics

Input data were modified by aggregation, affecting the regional mean and variability of the data. The magnitude of the modification is mainly due to two factors: i) the aggregation

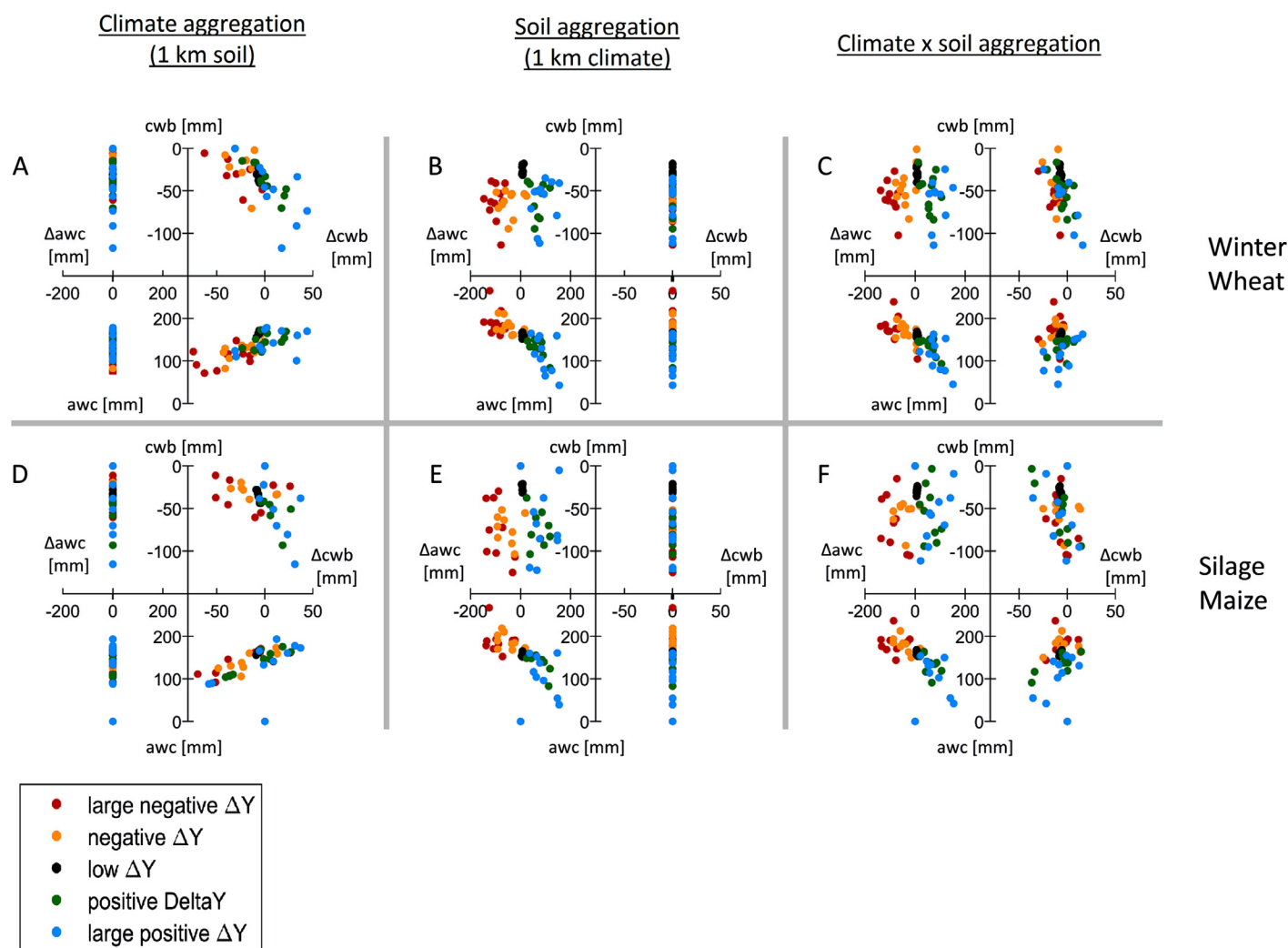


Fig 7. Differences in yield (ΔY) due to data aggregation as related to available water capacity (awc) and climatic water balance during the growing season (cwb) at 1 km resolution and the corresponding differences from coarser resolutions to 1 km resolution due to data aggregation (Δawc , Δcwb). Panels show results for winter wheat (A,B,C), silage maize yields (D,E,F) and three combinations of aggregated data: A,C: climate aggregated and 1 km soil resolution; B, D: soil aggregated and 1 km climate resolution; C, F: soil and climate aggregated simultaneously. Values shown are single model means for each of the five groups of aggregation effects (one scatter dot per model), grouped as follows. Large negative ΔY : $\Delta Y < \mu - 2\sigma$; negative ΔY : $\mu - 2\sigma < \Delta Y < \mu - \sigma$; low ΔY : $\mu - \sigma < \Delta Y < \mu + \sigma$; positive ΔY : $\mu + \sigma < \Delta Y < \mu + 2\sigma$; Large positive ΔY : $\Delta Y > \mu + 2\sigma$ where μ : mean of ΔY and σ : standard deviation of ΔY . See Table D in S1 Table for detailed values.

doi:10.1371/journal.pone.0151782.g007

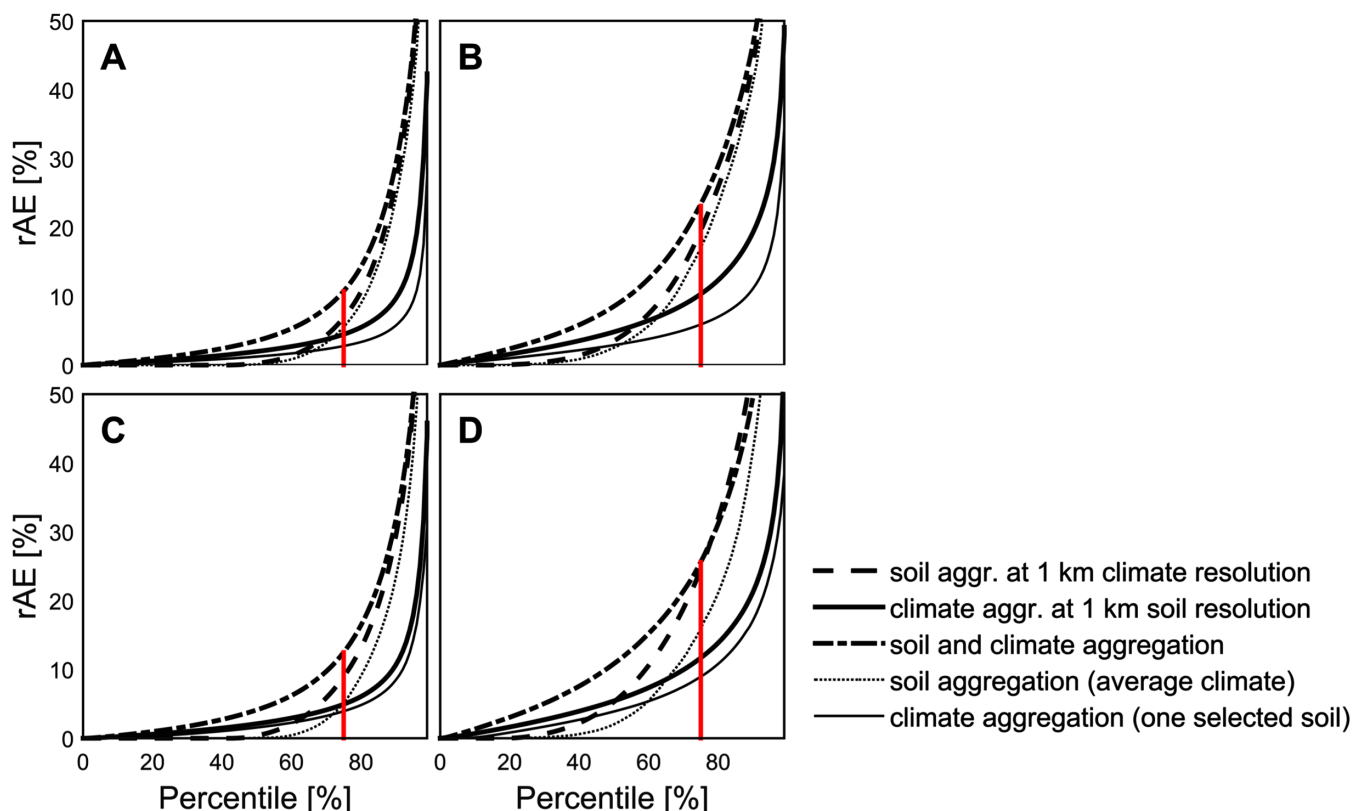


Fig 8. Relative absolute error (rAE) of yield under water-limited conditions from single cells as compared to 1 km resolution ($n = 34168$ cells \times 30 years \times 11 models = 11275440). (A) Winter wheat, aggregated data at 10 km resolution. (B) Winter wheat, aggregated data at 100 km resolution. (C) Silage maize, aggregated data at 10 km resolution. (D) Silage maize, aggregated data at 100 km resolution. The red line highlights the 75% percentile of combined soil and climate aggregation.

doi:10.1371/journal.pone.0151782.g008

method and ii) the spatial variability of the data at the finest resolution. Exemplarily, it was shown in the present study, how the spatial variability of climate time series decreases with spatial averaging. This is in agreement with [11] who also showed the spatial dependency of the climate data (variogram). Considering the aggregation method *averaging* and the larger spatial kriging range of most climate variables as compared to soil data, aggregated climate data was expected to exhibit continuous (“smooth”) aggregation effects with coarser resolutions. In contrast, soil data aggregation led to abrupt changes in the data statistics due to the selected method. These changes are the result of the changing mixture of soil types in space with aggregation which leads to large differences when compared to the highest resolution. Furthermore, these abrupt changes are not monotonously increasing or decreasing and the aggregated data thus depend highly on the resolution. However, effects on simulated yields may be rather small regarding regional statistics. This is in agreement with [25], who found decreasing extremes but similar 5th and 95th percentiles in soil water storage capacity after aggregating soil types from 1 to 100 km resolution by area majority for Germany. Thus, the prevailing spatial variability of the soil data in this study could be representative for a larger fraction of soils in Central Europe. Large-scale crop modelling studies may therefore be robust to such data artefacts on average, but prone to errors at higher resolutions, especially when using spatially aggregated soil data. For instance, the selected resolution in such studies may lead to different directions and magnitude of aggregation effects of input data in different regions. Therefore, new

concepts are needed to integrate these effects into the assessment of uncertainty in large-scale studies when making statements on sub-regions or spatial patterns.

Relationship of simulated water-limited and potential yield in aggregation effects

Soil and climate data aggregation effects on yield may be larger in extreme situations. For instance, [11] hypothesized that large positive effects on yield would be observed best in situations with high abiotic stress. In contrast, large negative effects would be observed best under near-optimum conditions. Since the modelled yields include simulations from soils and climatic periods with non-optimum conditions as well as homogeneous management, it is questionable how much abiotic stress prevailed in the present simulations. Simulated yields of 7.2 to 8.1 t ha⁻¹ (winter wheat) and 15.2 to 16.0 t ha⁻¹ (silage maize) in the median of models were similar to observed yields of 7.2 and 14.3 t ha⁻¹ [11]. With an annual precipitation larger than 800 mm (regional median), a high degree of drought stress was not expected [11–12]. For the same investigated region, [11–12] simulated yields with a similar model ensemble for potential production situations, where crop growth was only limited by temperature and radiation input data. The yield gaps obtained between potential and water-limited conditions were 4.7 and 6.9% for winter wheat and maize respectively. In these studies, a single soil profile of high awc was applied for the entire region, equivalent to the combination of a single selected soil vs. climate data aggregation applied in the present study. Here we show regional mean yields of 0.9 t ha⁻¹ (winter wheat) and 0.8 t ha⁻¹ (silage maize) lower than for the mentioned single soil profile. With this single soil profile, [11–12] obtained similar yields as presented here (differences ≤ 0.2 t ha⁻¹). Therefore, abiotic stress due to water limitation can be approximated by relating the potential yields of the mentioned studies to the water-limited yields of the present study. As a result the average yield reduction of all years due to water deficits in the present study turned out to be 15.1 and 11.5% for winter wheat and silage maize, respectively. Since water scarcity is mainly restricted to sub-regions of low awc and sporadically low cwb, larger rAE and rMAE could be expected for the present study, in particular due to differences in the soil input data.

Effects of data aggregation on simulated yields

Interpretation of aggregation effects and comparison to other sources of yield variability. Using aggregated climate and/or soil model input data lead to errors in simulated water-limited yields of winter wheat and silage maize. In the present study, regional yield biases of up to $\pm 3.1\%$ (ensemble median) were in the range of published studies (Table 1). Furthermore, these ensemble median biases were relatively small compared to the inter-annual yield variability or variability of yields across crop models. However, two exceptions of considerably larger aggregation effects could be recognised. Firstly, few models showed considerably larger biases of up to 42.7%. In addition, model biases partially ranged from negative to positive values. For instance, model biases for silage maize yield ranged from -3.1 to +6.1% with climate input data aggregation in combination with soil data of 1 km resolution. Since the median of biases across models was +2.3% in this case, the model ensemble median seems to be a rather poor indicator of positive or negative aggregation effects. Secondly, the rMAE was considerably larger than yield biases, partially exceeding the inter-annual or inter-model variability in yields. Therefore, aggregation errors can be a dominating source of error when regarding the rMAE calculated from single years. In contrast, these errors cancel out in a long-term average. Finally, the rMAE calculated from yields averaged over the years (data not shown) compared well to the relative average absolute deviation (rAAD) reported by [12] for climate data aggregation in combination with homogeneous soil input data. This rMAE is lower than the rMAE reported here as

calculated from single years. Thus, caution should be taken when interpreting results as the different error statistics report larger errors i) when calculated from single years and averaged afterwards as compared to calculations from averages of years and ii) for the rMAE than for the bias. This is illustrated by Figure G in [S1 Fig](#), where ensemble median of regional mean yields correlated well to yields at 1 km resolution ($R^2 > 0.95$), but single cell yields correlated poorly to yields at 1 km resolution ($R^2 \leq 0.31$) when aggregating both soil and climate input data.

Assessing regional soil and climate data aggregation effects. Aggregation effects have been investigated by a range of studies ([Table 1](#)). In order to relate these findings to the present study, the following steps were applied: 1) aggregation of one type of input data keeping other information constant at coarse resolution (e.g. one data input value per region); 2) aggregation of one type of input data keeping other information constant at high resolution (e.g. 1 km); 3) aggregating several types of input data simultaneously.

While step 1 aims at analysing the aggregation effect of one data type, it fails at assessing the interactions with the data type, which is not aggregated for realistic conditions. For instance, climate input data aggregation effects have been investigated by using a single soil profile across all simulations [[11–12](#), [15](#), [28](#)]. This is comparable to the single soil vs. climate aggregation of the present work. All of these studies agree well with the present work for low regional yield bias, being considerably lower as compared to inter-annual variability or model differences in yields. This suggests that climate data aggregation contributes less to the uncertainty in upscaling for yield simulations. However, these studies were carried out with temperate climates of positive annual water balances and typical agricultural soils of awc above 200 mm. Thus, these studies may not be suitable to portray the full variability or aggregation effects in other more-constraint climate environments. Using aggregated soils with average climate or single climate time series is analogue to this.

In contrast, the steps 2 and 3 given above account for the entire range of input data variability of a region. For instance, [[10](#)] aggregated different input data types of high resolution for an ecosystem model by averaging. By aggregating the different data types stepwise, they were able to quantify the contribution of each data type to the aggregation bias. But, except for single specific soil variables (e.g. awc), averaging soil information is usually not meaningful, since averaging may lead to strong inconsistencies among different variables. This is due to the nonlinear relations of the soil variables and the impact of averaging on texture classes, e.g. by averaging sand and clay fractions. And, when using different aggregation methods, effects on yield are not necessarily additive. Therefore, aggregation effects across soil and climate can only be assessed by comparing systematically outputs from simultaneously aggregated inputs with outputs from separately aggregated inputs (keeping one input data type constant) and simulating a large range of possible input data combinations (large spatial variability) for various resolutions. To do so [[24](#)] used CLIMCROP to simulate winter wheat yields with “fine resolution” soil data (Danish soil survey) as well as with one representative soil per county, keeping climate data at 1 km resolution. A maximum bias of 13.9% of simulated to observed yields was reported. Although this finding is not directly comparable to the present study due to the different resolutions, grids and underlying data used to construct the grids, it is similar to biases found here.

It is also possible to derive the optimum data resolution by knowing the regional yield bias. Therefore, [[24](#)] concluded that a minimum soil resolution of 10 km is needed to upscale crop simulations for Denmark. Similarly, [[23](#)] found the best agreement of observed and simulated yields with climate and soil data at approximately 1° resolution (approx. 100 km). In the present study, yields were compared to yields of the highest resolution instead of regional statistics. This was done to calculate the direct data aggregation effect, ruling out artefacts in aggregating

yield statistics. However, we found ensemble median biases of less than 3% across all resolutions and data type combinations as well as maximum single model biases of less than 4% for up to 25 km resolution and less than 8% for up to 50 km resolution. Depending on the aim, this suggests keeping simulation resolutions roughly in the range of up to 25 or 50 km when the model's sensitivity to soil data aggregation is not known. By contrast, findings from climate data aggregation studies indicated a reasonable threshold at 100 km resolution for regions in Central Europe [11–12, 30]. Additionally, [25] suggested a resolution of 100 km as reasonable for aggregating soil and climate data for Germany. In general, all of these findings are comparable with the median bias across models of the current work. This emphasizes the dominant influence of spatial soil variability on spatial yield variability and eventually aggregation effects in Europe, being similar to findings by [22]. [28] suggested to adapt the resolution of climate input data to the spatial heterogeneity of relevant variables, e.g. topography. With this approach, they were able to reduce the root mean squared error (RMSE) in winter wheat yield when aggregating climate data. While promising, the performance of the approach remains to be assessed for other input data types and aggregation methods.

Patterns in soil and climate aggregation effects. The presented rMAE show large differences across data types and aggregation methods, which are not captured by the regional yield bias when using a model ensemble median. It would thus be intriguing to test, how much single models were influenced by either soil or climate data aggregation. While a detailed assessment of this aspect is beyond the scope of this work, the presented scatter of model aggregation effects, grouped by their magnitude, can serve as an initial indicator. We have shown for all models that the direction and magnitude of the aggregation effect depends on the yield and awc at 1 km resolution, and the change in awc when aggregating soil or soil and climate. This relationship may differ for regions with less precipitation. Still, if yields can be approximated by few variables, mainly soil related variables as in the present study, we hypothesize that the aggregation error can thus be approximated. By knowing the soil properties and their distribution, changes in soil due to aggregation by majority are known. Thus, the direction and magnitude of the aggregation effect may be estimated before conducting large-scale studies. If a model response (e.g. simulated yield) on each soil type is known and the model response to climate is low or climate is spatially homogeneous, then aggregation effects could be estimated by calculating the expected differences from changing spatial soil type composition. However, when extrapolating for different regions, this is restricted by the accuracy and robustness in relating a models yield to the given soil conditions. Other approaches could follow [28], possibly incorporating further or different variables as proxies.

Aggregation effects of soil and climate data were not additive in the present study. However, we showed that combined soil and climate data aggregation can double the aggregation error in parts of the aggregation error distribution. This was due to the different methods of aggregation. Climate data averaging led to a larger proportion of low aggregation errors already at 10 km resolution, while soil aggregation led to a larger fraction of grid cells and years with no error as well as in the fraction of largest errors. Additionally, the maximum increase in the aggregation error due to combined climate and soil data aggregation was specifically observed between the points of inflexion of aggregation errors due to climate aggregation and those due to soil data aggregation. This additional aggregation error could thus possibly serve as an indicator of crop model sensitivity to errors from aggregated soil and climate data.

Limitations and link to aggregation effect assessment

Firstly, two aggregation methods were applied on two data types. Therefore, climate data averaging effects are not directly comparable to soil data aggregation by majority. For a complete

analysis of the variances in aggregation effects, it would thus be necessary to average soil data and select climate time series by majority. However, due to required consistency across variables, soil data averaging was not feasible. Similarly, selecting climate time series by majority is not feasible without previous classification, since all climate time series differ. The presented results therefore primarily demonstrate possible ranges and patterns of aggregation effects. While the aggregation methods used in this work are most common, large-scale studies use a larger range of aggregation and related scaling methods [13–14]. Furthermore, aggregation of other data types than used here may affect the aggregation differently. For instance, a constant management (sowing, fertilization) was used in the present study. However, management may be based on decision rules. Using aggregated management based on these rules at 1 km resolution or applying them directly at coarser resolution (e.g. aggregated climate) could lead to aggregation errors. This remains to be confirmed.

Secondly, crop models were calibrated in order to represent regional long-term mean yields over the entire study area. Thus, models which were tuned to fit observed yields in more detail (e.g. in single years or counties) are possibly less robust to input data aggregation as presented here. Finally, the effects of spatial data aggregation on temporal in- and output patterns are unknown. For instance, temporal scale-invariant time series parameters [77] may be affected by spatial aggregation. As this is potentially relevant for impact assessment, this remains to be investigated.

Conclusions

We conclude that, when simulating regional water-limited average yields in a temperate humid region, most models are little affected by aggregating soil and/or climate data up to 100 km resolution. However, some models showed considerably larger biases in the range or larger than the inter-annual yield variability. Consequently, models need to be assessed individually for their robustness to climate and soil input data aggregation when simulating regional yields. Furthermore, absolute errors (rMAE) of most models using aggregated soil data are in the range or larger than the inter-annual yield variability or differences between models. Being even larger than aggregation effects from averaged climate data, aggregating soils can thus be a dominant source of uncertainty when assessing spatial yield patterns of heterogeneous regions. Furthermore, simultaneous use of aggregated climate and soil data is likely to increase these aggregation effects even further. However, aggregation effects in yield show distinct patterns depending on the type of data being combined. Large negative aggregation effects were found in areas with soils characterized by high available water holding capacity and large positive aggregation effects in areas with soils of predominantly low available water holding capacity. Considering the regional precipitation pattern, this indicates that the direction and magnitude of aggregation effects may be estimated from a limited number of soil variables.

Supporting Information

S1 Dataset. Minimal dataset with underlying data of presented results.
(ZIP)

S1 Fig. Supporting Figures A-G.
(PDF)

S1 File. Equations.
(DOCX)

S1 Table. Supporting Tables A-D.
(PDF)

Acknowledgments

This work was financially supported by the German Federal Ministry of Food and Agriculture (BMEL) through the Federal Office for Agriculture and Food (BLE), (2851ERA01J). FT and RPR were supported by FACCE MACSUR (3200009600) through the Finnish Ministry of Agriculture and Forestry (MMM). EC, HE and EL were supported by The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (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) and thank professor P-E Jansson (Royal Institute of Technology, Stockholm) for support. JC, HR and DW thank the INRA ACCAF metaprogramm for funding and Eric Casellas from UR MIAT INRA for support. CB was funded by the Helmholtz project “REKLIM—Regional Climate Change”. CK was funded by the HGF Alliance “Remote Sensing and Earth System Dynamics” (EDA). FH was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) under the Grant FOR1695. FE and SS acknowledge support by the German Science Foundation (project EW 119/5-1). HH, GZ, SS, TG and FE thank Andreas Enders and Gunther Krauss (INRES, University of Bonn) for support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

Conceived and designed the experiments: HH JC RD HE TG BG KCK MK EL CN HR DW LW FE EP PPR RPR JY SA MB. Performed the experiments: HH GZ CB JC EC LD BG FH BTK KCK CK MK MM HR XS FT ET GT LW. Analyzed the data: HH GZ SS. Wrote the paper: HH.

References

1. Folberth C, Yang H, Wang X, Abbaspour KC. 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*. 2012; 235: 8–18.
2. Rosenzweig C, Elliott J, Deryng D, Ruane AC, Müller C, Arneth A, et al. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *P Natl Acad Sci USA*. 2014; 111, 3268–3273.
3. Van Ittersum M, Leffelaar P, Van Keulen H, Kropff M, Bastiaans L, Goudriaan J. On approaches and applications of the Wageningen crop models. *Eur J Agron*. 2003; 18: 201–234.
4. Hansen JW, Challinor A, Ines A, Wheeler T, Moron V. Translating climate forecasts into agricultural terms: advances and challenges. *Clim Res*. 2006; 33: 27–41.
5. Meentemeyer V. Geographical perspectives of space, time, and scale. *Landscape Ecol*. 1989; 3: 163–173.
6. Hansen JW, Jones JW. Scaling-up crop models for climate variability applications. *Agr Syst*. 2000; 65: 43–72.
7. Baron C, Sultan B, Balme M, Sarr B, Traore S, Lebel T, et al. From GCM grid cell to agricultural plot: scale issues affecting modelling of climate impact. *Philos T R Soc B*. 2005; 360: 2095–2108.
8. Cale WG, Oneill RV, Gardner RH. Aggregation error in non-linear ecological models. *Journal of Theoretical Biology*. 1983; 100: 539–550.
9. Rastetter EB, King AW, Cosby BJ, Hornberger GM, Oneill RV, Hobbie JE. Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. *Ecol Appl*. 1992; 2: 55–70.
10. Pierce LL, Running SW. The effects of aggregating subgrid land-surface variation on large-scale estimates of net primary production. *Landscape Ecology*. 1995; 10: 239–253.
11. Hoffmann H, Zhao G, van Bussel LGJ, Enders A, Specka X, Sosa C, et al. Variability of aggregation effects of climate data on regional yield simulation by crop models. *Clim Res*. 2015; 65: 53–69.
12. Zhao G, Hoffmann H, Van Bussel LGJ, Enders A, Specka X, Sosa C, et al. Effect of weather data aggregation on regional crop simulation for different crops, production conditions, and response variables. *Clim Res*. 2015; 65: 141–157.

13. Ewert F, Van Ittersum MK, Heckelet T, Therond O, Bezlepina I, Andersen E. Scale changes and model linking methods for integrated assessment of agri-environmental systems. *Agr Ecosyst Environ*. 2011; 142: 6–17.
14. Ewert F, Van Bussel LGJ, Zhao G, Hoffmann H, Gaiser T, Specka X, et al. Uncertainties in Scaling up Crop Models for Large Area Climate-change Impact Assessments. In Rosenzweig C. and Hillel D. (Eds.). *Handbook of Climate Change and Agroecosystems: The Agricultural Model Intercomparison and Improvement Project (AgMIP)*. World Scientific Publishing Company. 2015; pp. 262–277.
15. Angulo C, Rötter R, Trnka M, Pirttioja N, Gaiser T, Hlavinka P, Ewert F. Characteristic 'fingerprints' of crop model responses data at different spatial resolutions to weather input. *Eur J Agron*. 2013; 49: 104–114.
16. Angulo C, Gaiser T, Rötter RP, Børgesen CD, Hlavinka P, Trnka M, Ewert F. 'Fingerprints' of four crop models as affected by soil input data aggregation. *Eur J Agron*. 2014; 61: 35–48.
17. Nendel C, Wieland R, Mirschel W, Specka X, Guddat C, Kersebaum KC. Simulating regional winter wheat yields using input data of different spatial resolution. *Field Crop Res*. 2013; 145: 67–77.
18. Janssen S, Andersen E, Athanasiadis IN, Van Ittersum MK. A database for integrated assessment of European agricultural systems. *Environ Sci Policy*. 2009; 12: 573–587.
19. Zhao G, Bryan BA, King D, Luo Z, Wang E, Song X, Yu Q. Impact of agricultural management practices on soil organic carbon: simulation of Australian wheat systems. *Glob Change Biol*. 2013; 19: 1585–1597.
20. Therond O, Hengsdijk H, Casellas E, Wallach D, Adam M, Belhouchette H, et al. Using a cropping system model at regional scale: Low-data approaches for crop management information and model calibration. *Agr Ecosyst Environ*. 2011; 142: 85–94.
21. Webber H, Zhao G, Wolf J, Britz W, De Vries W, Gaiser T et al. Climate change impacts on European crop yields: do we need to consider nitrogen limitation? *Eur J Agron*. 2015; 71: 123–134.
22. Wassenaar T, Lagacherie P, Legros JP, Rounsevell MDA. Modelling wheat yield responses to soil and climate variability at the regional scale. *Clim Res*. 1999; 11: 209–220.
23. Easterling WE, Weiss A, Hays CJ, Mearns LO. Spatial scales of climate information for simulating wheat and maize productivity: the case of the US Great Plains. *Agr Forest Meteorol*. 1998; 90: 51–63.
24. Olesen JE, Bøcher PK, Jensen T. Comparison of scales of climate and soil data for aggregating simulated yields of winter wheat in Denmark. *Agr Ecosyst Environ*. 2000; 82: 213–228.
25. Eyshi Rezaei E, Siebert S, Ewert F. Impact of data resolution on heat and drought stress simulated for winter wheat in Germany. *Eur J Agron*. 2015; 65: 69–82.
26. De Wit AJW, Boogaard HL, Van Diepen CA. Spatial resolution of precipitation and radiation: The effect on regional crop yield forecasts. *Agr Forest Meteorol*. 2005; 135: 156–168.
27. Ershadi A, McCabe MF, Evans JP, Walker JP. Effects of spatial aggregation on the multi-scale estimation of evapotranspiration. *Remote Sens Environ*. 2013; 131: 51–62.
28. Zhao G, Siebert S, Enders A, Rezaei EE, Yan C, Ewert F. Demand for multi-scale weather data for regional crop modelling. *Agr For Meteorol*. 2015; 200: 156–171.
29. Kersebaum KC & Wenkel KO. Modelling water and nitrogen dynamics at three different spatial scales—influence of different data aggregation levels on simulation results. *Nutr Cycl Agroecosys*. 1998; 50: 313–319.
30. Van Bussel LGJ, Ewert F, Leffelaar PA. Effects of data aggregation on simulations of crop phenology. *Agr Ecosyst Environ*. 2011; 142: 75–84.
31. Cologne district council. Nordrhein-Westfalen in Zahlen und Geodaten. 2013. URL: www.bezreg-koeln.nrw.de/brk_internet/publikationen/abteilung07/pub_geobasis_nrw.pdf (last fetch: 15.12.2015)
32. FAO 2015. World reference base for soil resources 2014. International soil classification system for naming soils and creating legends for soil maps. Update 2015. Food and Agricultural Organization of the United Nations. World Soil Resources Reports 106. ISSN 0532-0488.
33. Wendling U. Berechnung der Gras-Referenzverdunstung mit der FAO Penman-Monteith-Beziehung. *Wasserwirtschaft*. 1995; 85: 602–604.
34. 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 (accessed: January 27, 2015).
35. AG-Boden. Ad hoc Arbeitsgruppe Boden: Bodenkundliche Kartieranleitung, 5th ed. Federal Institute for Geosciences and Natural Resources in cooperation with the Federal Geological Services, Hannover; 2005. 438 pp.

36. LANUV. North Rhine-Westphalia State Agency for Nature, Environment and Consumer Protection. Unit 33, soil conservation, Recklinghausen, Germany. 2014. <http://www.lanuv.nrw.de> (accessed: January 25, 2015).
37. Poesen J, Lavee H. Rock fragments in top soils: significance and processes. *Catena*. 1994; 23: 1–18.
38. Torri D, Poesen J, Monaci F, Busoni E. Rock fragment content and fine soil bulk density. *Catena*. 1994; 23: 65–71.
39. Hoffmann H, Enders A, Zhao G, Siebert S, Gaiser T, Ewert F, 2016. Climate and soil input data aggregation effects in crop models. URL: 10.7910/DVN/C0J5BB, Harvard Dataverse.
40. Evans LT, Fisher RA. Yield potential: Its definition, measurement, and significance. *Crop Science*. 1999; 39: 1544–1551.
41. Van Ittersum MK, Rabbinge R. Concepts in production ecology for analysis and quantification of agricultural input-output combinations. *Field Crop Res*. 1997; 52: 197–208.
42. Statistische Ämter des Bundes und der Länder. Regionaldatenbank Deutschland. 2013. URL: <https://www.regionalstatistik.de/genesis/online/logon>; accessed 08/05/2014.
43. Herbst M, Hellebrand HJ, Bauer J, Huisman JA, Šimůnek J, Weihermüller L, et al. Multiyear heterotrophic soil respiration: Evaluation of a coupled CO₂ transport and carbon turnover model. *Ecological Modelling*. 2008; 214: 271–28.
44. Asseng S, Keating BA, Fillery IRP, Gregory PJ, Bowden JW, Turner NC, Palta JA, Abrecht DG. Performance of the APSIM-wheat model in Western Australia. *Field Crop Res*. 1998; 57: 163–179.
45. Asseng S, Jamieson PD, Kimball B, Pinter P, Sayre K, Bowden JW, Howden SM. Simulated wheat growth affected by rising temperature, increased water deficit and elevated atmospheric CO₂. *Field Crop Res*. 2004; 85: 85–102.
46. Keating BA, Carberry PS, Hammer GL, Probert ME, Robertson MJ, Holzworth DP, et al. An overview of APSIM, a model designed for farming systems simulation. *Eur J Agron*. 2003; 18: 267–288.
47. Holzworth DP, Huth NI, De Voil PG, Zurcher EJ, Herrmann NI, McLean G, et al. APSIM—Evolution towards a new generation of agricultural systems simulation. *Environ Model Softw*. 2014; 62: 327–350.
48. Lugato E, Berti A, Giardini L. Soil organic carbon (SOC) dynamics with and without residue incorporation in relation to different nitrogen fertilisation rates. *Geoderma*. 2006; 135: 315–321.
49. Parton WJ, McKeown B, Kirchner V, Ojima D. CENTURY Users Manual. Colorado State University, NREL Publication, Fort Collins, Colorado, USA. 1992.
50. Jansson PE. Coup Model: model use, calibration and validation. *Transactions of the ASABE (American Society of Agricultural and Biological Engineers)*. 2012; 55: 1335–1344.
51. Conrad Y, Fohrer N. Modelling of nitrogen leaching under complex winter wheat and red clover crop rotation on a drained agricultural field. *Physics and Chemistry of the Earth*. 2009; 34: 530–540.
52. Jansson PE, Karlberg L. Coupled heat and mass transfer model for soil-plant-atmosphere systems, Royal Institute of Technology, Department of Civil and Environmental Engineering, Stockholm, Sweden. 2004; 435 pp. URL: <http://www2.lwr.kth.se/Vara%20Datorprogram/CoupModel/NetHelp/default.htm> (last fetch: 15.12.2014)
53. Stöckle CO, Donatelli M, Nelson R. CropSyst, a cropping systems simulation model. *Eur J Agron*. 2003; 18: 289–307.
54. Del Grosso SJ, Parton WJ, Mosier AR, Hartman MD, Brenner J, Ojima DS, Schimel DS. Simulated interaction of carbon dynamics and nitrogen trace gas fluxes using the DAYCENT model. In: Schaffer M (Ed.), *Modeling Carbon and Nitrogen Dynamics for Soil Management*. CRC Press, Boca Raton, Florida, USA, 2001; pp. 303–332.
55. Del Grosso SJ, Parton W, Mosier AR, Walsh MK, Ojima D, Thornton PE. DAYCENT national scale simulations of N₂O emissions from cropped soils in the USA. *J. Environ. Qual*. 2006; 35: 1451–1460. PMID: 16825465
56. Parton WJ, Holland EA, Del Grosso SJ, Hartman MD, Martin RE, Mosier AR, et al. Generalized model for NO_x and N₂O emissions from soils. *J. Geophys. Res*. 2001; 106: 17403–17420.
57. Yeluripati JB, Van Oijen M, Wattenbach M, Neftel A, Ammann A, Parton WJ, Smith P. Bayesian calibration as a tool for initialising the carbon pools of dynamic soil models. *Soil Biol Biochem*. 2009; 41: 2579–2583.
58. Priesack E, Gayler S, Hartmann HP. The impact of crop growth sub-model choice on simulated water and nitrogen balances. *Nutr Cycl Agroecosys*. 2006; 75: 1–408.
59. Wang E, Engel T. SPASS: a generic process-oriented crop model with versatile windows interfaces. *Environ Modell Softw*. 2000; 15: 179–188.
60. Williams JR. The EPIC model. In: Singh V.P. (Ed.), *Computer models of watershed hydrology*. Water resources publications, Highland Ranch CO. 1995; pp. 909–1000.

61. Kersebaum KC. Modelling nitrogen dynamics in soil-crop systems with HERMES. *Nutr Cycl Agroecosys*. 2007; 77: 39–52.
62. Kersebaum KC. Special features of the HERMES model and additional procedures for parameterization, calibration, validation, and applications In: Ahuja L.R. and Ma L. (ed.): *Advances in Agr Syst Modeling Series*. 2011; 2: 65–94.
63. Tao F, Yokozawa M, Zhang Z. Modeling the Impacts of Weather and Climate Variability on Crop Productivity over a Large Area: A New Process-based Model Development, Optimization, and Uncertainties Analysis. *Agr Forest Meteorol*. 2009; 149: 831–850.
64. Tao F, Zhao Z. Climate change, wheat productivity and water use in the North China Plain: A new super-ensemble-based probabilistic projection. *Agr Forest Meteorol*. 2013; 170: 146–166.
65. Nendel C, Berg M, Kersebaum KC, Mirschel W, Specka X, Wegehenkel M, et al. The MONICA model: Testing predictability for crop growth, soil moisture and nitrogen dynamics. *Ecol Model*. 2011. 222: 1614–1625.
66. Specka X, Nendel C, Wieland R. Analysing the parameter sensitivity of the agro-ecosystem model MONICA for different crops. *Eur J Agron*. 2015; 71: 73–87.
67. Shibu ME, Leffelaar PA, van Keulen H, Aggarwal PK. LINTUL3, a simulation model for nitrogen-limited situations: Application to rice. *Eur J Agron*. 2010; 32: 255–271.
68. Bergez JE, Chabrier P, Gary C, Jeuffroy MH, Makowski D, Quesnel G, et al. An open platform to build, evaluate and simulate integrated models of farming and agro-ecosystems. *Environ Modell Softw*. 2013; 39: 39–49.
69. Brisson N, Mary B, Ripoche D, Jeuffroy MH, Ruget F, Nicoulaud B, et al. STICS: a generic model for the simulation of crops and their water and nitrogen balances. 1. Theory and parameterization applied to wheat and corn. *Agronomie*. 1998; 18: 311–346.
70. Brisson N, Launay M, Mary B, Beaudoin N. Conceptual basis, formalisations and parameterization of the STICS crop model, Quae. ed. 2009. ISBN 978-2-7592-0169-3.
71. Taylor KE. Summarizing multiple aspects of model performance in a single diagram. *J Geophys Res-Atmos*. 2001; 106: 7183–7192.
72. Hoffmann H, Rath T. Future bloom and blossom frost risk for *Malus domestica* considering climate model and impact model uncertainties. *PLoS ONE*. 2013; doi: [10.1371/journal.pone.0075033](https://doi.org/10.1371/journal.pone.0075033)
73. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. *Chemometr Intell Lab Syst*. 2001; 58: 109–130.
74. LANUV. North Rhine-Westphalia State Agency for Nature, Environment and Consumer Protection. Klimaatlas Nordrhein-Westfalen. Klimatische Wasserbilanz. <http://www.klimaatlas.nrw.de/site/nav2/Niederschlag.aspx?P=3&M=2> (accessed: October 3, 2015).
75. LANUV. North Rhine-Westphalia State Agency for Nature, Environment and Consumer Protection. Klimafolgenmonitoring. 6—Klimatische Wasserbilanz. <http://www.lanuv.nrw.de/kfm-indikatoren/index.php?indikator=45&mode=indi&aufzu=1> (accessed: October 3, 2015).
76. Krzywinski M. Visualizing samples with box plots. *Nature Methods*. 2014; 11: 119–120. PMID: [24645192](https://pubmed.ncbi.nlm.nih.gov/24645192/)
77. Baranowski P, Krzyszczyk J, Slawinski C, Hoffmann H, Kozyra J, Nieróbca A, et al. Multifractal analysis of meteorological time series to assess climate impacts. *Clim Res*. 2015; 65: 39–52.