DOI: https://doi.org/10.1016/j.agrformet.2015.10.005

Statistical regression models for assessing weather impacts on crop yields – A validation study for winter wheat and silage maize in Germany

Christoph Gornott[1*] and Frank Wechsung[1]

[1] Potsdam Institute for Climate Impact Research (PIK)

[*] Corresponding author

Abstract

For agriculture in Germany and generally all around the world, yield variability due to uncertain weather conditions represents an increasing production risk. Regional assessments of future yield changes can help farmers to cope with this risk. For Germany's two most important crops winter wheat (*Triticum aestivum* L.) and silage maize (*Zea mays* L.), we investigate three regression models estimating relative weather impacts on relative crop yield changes: the separate time series model (STSM), the panel data model (PDM) and the random coefficient model (RCM). These regression models use the Cobb–Douglas function to capture weather and non-weather impacts on yields (e.g. changing prices or inventory management). The yield influencing weather impacts contain the potential growth and stress factors during vegetative and reproductive plant development. The models are estimated and validated at the county scale. To improve the robustness and goodness of fit, the models are aggregated at the scale of German federal states, river basins and at the national scale. The observed yield changes are satisfactorily reproduced by all models for all aggregated scales (measured by the Nash–Sutcliffe efficiency (NSE)). According to their NSE values, the methodically simple STSMs reproduce extreme yield changes better (0.85) than the RCMs (0.79) and PDMs (0.72) at the national scale. This order can be also found across all scales when considering the models' goodness of fit. Generally, spatial aggregation increases the goodness of fit by +0.16 for federal states and river basins and by +0.29 for entire Germany compared to the county scale. The mean NSE increase is lowest for STSMs (+0.11), followed by RCMs (+0.13) and PDMs (+0.25) for federal states and river basins, which is opposite to the goodness of fit order. The model parameters show clear spatial patterns, which reflect regional differences of weather and soil. Within its methodological limits, our approach can directly be combined with the output of climate models and is suitable for assessing short- and medium-term yield effects for the current agronomic practice. It requires neither bias correction of the climate variables nor explicit modeling of crop yield trends.

Keywords:

Statistical crop yield models, Weather impacts, Yield changes, Winter wheat, Silage maize, Germany

Introduction

Statistical crop models for yield assessments

Crop yield assessments for upcoming climate anomalies or altered weather conditions are of general interest for farmers, traders (e.g. grain mills, retailers), insurance companies, and policy makers. Statistical models (Ray et al., 2015; Iizumi et al., 2013; Mueller et al., 2012; Roberts et al., 2012; Schlenker and Roberts, 2009) and process based models (Asseng et al., 2013; Angulo et al., 2013; Palosuo et al., 2011) are model types for such assessments. Both model types are parametrized for past weather records. For future projections, they need weather records from climate simulation models. These climate models very often require a bias correction of the simulated output before using them for yield projections (Lobell, 2013).

Process based crop models may not include all weather related effects on crop yields. There are many yield effects, which simply cannot be captured in process based models, because of limited spatial information about these effects. Examples are weather-triggered effects on agronomic adaptation (irrigation, crop varieties, agronomic techniques) or on pests, weeds, and diseases (Mueller et al., 2012). These weather-triggered effects can be collinear with the weather variables. Since crop yields also contain weather-triggered effects, statistical crop yield models estimate in their parameter values not only the sole, but also the triggered effect of the weather variable. Process based models do not capture these weather-triggered effects as long as they are not explicitly embedded in the models (Estes et al., 2013; Lobell and Burke, 2010). In the assessment of farm level yield effects, this is an important disadvantage of process based models in comparison to statistical models.

Statistical yield models also allow relating inter-annual yield and yield factor changes (i.e. first order temporal ratios) instead of absolute values to each other (You et al., 2009; Lobell, 2007; Lobell and Asner, 2003). Considering changes instead of absolute values eliminates the trend of the variables and it allows neglecting systematic biases for exogenous variables for example when using simulated climate data from circulation models (Lobell, 2013). However, the neglected absolute level by using changes ignores a possible level dependency of yield and climate conditions. This limits the suitability to climate change assessments for changes within the range of recent climate variability. For yield projections beyond the yield variability of the dataset used for model estimation, process based models might be more appropriate (Rötter et al., 2011). At least, process based models should complement the statistical assessments under such circumstances.

The impact of weather on crop yields can be subdivided into two variable groups: variables that primarily determine potential growth and those that can be related to stress influences. The distinction is not disjunctive, overlaps might exist. We focus on the main influences that can contribute to a statistical explanation of the yield variability. The potential yield is determined mainly by the incoming solar radiation (Monteith, 1977; Long et al., 2006). The best usage of this incoming solar radiation requires an optimal mix of agronomic measures to establish the crop, to supply the necessary nutrients and water, and to keep biotic stress factors under control. Any divergence from this optimal mix will result in stress that reduces the potential yield. For these potential stress factors, we distinguish two groups: weather and management driven stress factors.

Among all possible weather driven stress factors, we hypothesize water stress as the most relevant stress factor for German winter wheat and silage maize yields (Wessolek and Asseng, 2006; Kersebaum and Nendel, 2014; Wolf and Diepen, 1994). Other possible influences, like temperature stress, might also exist in single years (Rötter and van de Geijn, 1999; Lobell et al., 2013), but are less generally associable with German weather conditions. Management driven stress factors, like the crop variety, fertilizer, plant protection, and machinery, are reflected in the mean yield level and the yield trend. However, there are also economic conditions, e.g., statutory set-aside quotas or renewable energy subsidies for biogas and biodiesel, which influence the annual yield variability (Krause, 2008; Bakker et al., 2005). We use the fertilizer price and the acreage of the respective crops as proxy variables to control the economic yield impacts in the models. The fertilizer price represents the varying profitability of production factor inputs (e.g. seeds, plant protection, fuel, and fertilizer) and may directly affect the yield variability. The acreage of winter wheat and silage maize represents changes in the Common Agricultural Policy (CAP) of the European Union. An expanded acreage might generally suppress the yield level of both crops due to the inclusion of marginal productive land.

Modeling approach

In our approach, we follow the modeling concept introduced by Wechsung et al. (2008) and the validation scheme of Gornott and Wechsung (2015), who expanded the concept by two other statistical approaches. A level neutralizing transformation is applied for all variables, i.e., the crop yield, the weather-related and the non- weather-related variables. We utilize first order ratios $y_t' = \frac{y_t}{y_{t-1}}$ and $x_t' = \frac{x_t}{x_{t-1}}$, for the years $t = 2, \dots, M$ of the endogenous variable crop yield y_t and the exogenous weather and non- weather variables x_t. As functional form, we use the Cobb–Douglas function analogous to Oury (1965). The function is proven in both economic (You et al., 2009) and agronomic applications (Lee et al., 2013) and considers yield impacts arising from substitution and interaction between the exogenous variables. Due to the linearization of the Cobb–Douglas function, the first order ratios are transformed to logarithmic first order ratios of yields and yield-factors, hereafter expressed as yield and factor changes. These changes allow an intercomparison of the effects of different variables.

We test three alternative ways to incorporate the spatial heterogeneity of the relationships between yield changes and yield factor changes: by separately estimated time series models (STSMs), panel data models (PDMs), and random coefficient models (RCMs). All three approaches refer to a spatial dataset consisting of N discrete subunits and M years. In our case, the subunits are German counties within a federal state, river basin, or Germany as a whole. The methodically simple STSMs are estimated independently for the N subunits resulting in N parameter sets (Butler and Huybers, 2013; Lobell and Burke, 2010). In contrast, PDMs capture directly the temporal and spatial variability by one parameter set for all of the considered N subunits (You et al., 2009). RCMs can be ranked between PDMs and STSMs. They allow individual parameter variations per subunit and a parameter set for the entire unit (Reidsma et al., 2007). The results of the estimations will be presented and evaluated at two scales: the original spatial data scale, i.e., the German county yields, and the aggregated data scale, i.e. federal states, river basins, and entire Germany. Due to the aggregation, county- and farm-individual influences are largely averaged out, which might have biased the model results otherwise (Woodard and Garcia, 2008).

We restricted the temporal and spatial resolution of all variables to a division, which is accessible for climate simulations. The model results are evaluated at a larger scale than the estimation scale. Thus, we make explicit use of spatial aggregation effects. We test and apply the approach in respect to its possible suitability for fast impact assessment of seasonal- and medium-term projections (up to 30 years) from climate models. The approach is conducted for winter wheat and silage maize, because these are the major winter and summer annual crops in Germany.

Materials and methods

Data

We use a spatial dataset of German crop yields per county for winter wheat and silage maize from 1991 to 2010. The dataset is supplied by the Statistical Offices of the Federation and the Länder (2013b). Weather data are available for the same period from 1,218 German weather stations (DWD, 2011). The data are averaged per county to match the spatial resolution of the crop yield data. The total acreage of winter wheat and silage maize is taken from the datasets of the Statistical Offices of the Federation and the Länder (2013a) [1991–2008] and the Federal Statistical Office (2013) [2008–2010]. The fertilizer price index is published by the Statistical Offices of the Federation and the Länder (2013c). Ideally, all variables would be estimated at the county scale. However, the economic variables are only available on a national scale, so we applied the national values to all counties. A detailed description of the data is contained in the supplemental information (SI) S.1.

Model approach

Basic function

The Cobb–Douglas function is used as the basic function in all statistical models (Eq. 1). The function relates inter-annual changes of crop yield (y'_t) to J weather (x'_{jt}) and K economic variables (x'_{kt}). Statistical models often have the disadvantage that the parameter values are not easily accessible for an interpretation. In our approach, the parameter values of the Cobb–Douglas function are directly comparable per and across crops and spatial sites as relative yield effects by a relative increase of the exogenous variables (Wooldridge, 2013, p. 351-354).

$$y'_t = \beta_0 \prod_{j=1}^{J} \left(x'_{j\,t}\right)^{\beta_j} \prod_{k=1}^{K} (x'_{k\,t})^{\beta_k}, \text{with } j = 1, \dots, J \text{ and } k = 1, \dots, K \tag{1}$$

β_0 - mean annual changes of y'_t,

β_j, β_k - partial relative change of y' per unit change of x'_{jt} and x'_{kt}, respectively.

Regression models

The basic function (Eq. 1) can be linearized by logarithm. The variables are transformed to linear terms and the function is expanded by an error term u_t to become accessible for regression analysis. The spatial yield variability within Germany, German federal states and river basins is addressed using three alternative regression models: STSMs, PDMs, and RCMs.

STSMs (Eq. 2) separately consider the individual yield changes at the *N* subunits, by estimating independently a series of *N* models (Dielman, 1983).

$$\log\, y'_{it} = \log \beta_{i0} + \sum_{j=1}^{J} \beta_{ij} \log\, x'_{ijt} + \sum_{k=1}^{K} \beta_{ik} \log\, x'_{kt} + \log\, u'_{it}, \qquad i = 1, \dots, N \tag{2}$$

For our approach, the values of the x'_{kt} variables do not vary by the index *i* as all other variables, because county individual data of the economic variables are not available. Therefore only national values of those variables are related to the county individual crop yields.

Unlike STSMs, PDMs (Eq. 3) estimate directly one parameter set for all *N* subunits. The parameter values (β_j, β_k) do not vary among the *N* subunits as for STSMs. PDMs may still capture county individual, time-invariant effects (e.g. soil productivity and farm size effects) due to the normalizing effect of the county-wise first difference transformation. These effects, which are contained in the mean yield level, are eliminated by the first order transformation before model estimation (Wooldridge, 2013; Croissant and Millo, 2008). When this transformation is reversed after calculating absolute crop yields, then the spatial differences in the mean yield level re-appear.

$$\log\, y'_{it} = \log \beta_0 + \sum_{j=1}^{J} \beta_j \log\, x'_{ijt} + \sum_{k=1}^{K} \beta_k \log\, x'_{kt} + \log\, u'_{it} \tag{3}$$

RCMs (Eq. 4) may be ranked between STSMs and PDMs. They contain both one parameter set, which is valid for all subunits $(\beta_0, \beta_j, \beta_k)$ and county individual parameter variations (b_{i0}, b_{ij}, b_{ik}). The county-individual impact β_i results from $\beta_i = \beta + b_i$. Since the parameters β and b_i are dependent on each other, the model cannot be estimated by the ordinary least squares method (OLS). Instead, our RCMs are estimated by the restricted maximum likelihood method (REML) analogous to Reidsma et al. (2007).

$$\log\ y'_{it} = \log(\beta_0 + b_{i0}) + \sum_{j=1}^{J}(\beta_j + b_{ij}) \log\ x'_{ijt} + \sum_{k=1}^{K}(\beta_k + b_{ik}) \log\ x'_{kt} + \log\ u'_{it} \tag{4}$$

Aggregation of the model results

The estimated and measured N individual yield changes per county are averaged to the arithmetic mean (Eq. 5) for the aggregation scale (i.e. nation, federal states, and river basins) in hindsight. We did not aggregate the exogenous variables, because this would lead to information losses due to a reduced variability during the estimation.

$$\overline{\log y'_t} = N^{-1} \sum_{i=1}^{N} \log\ y'_{it} \tag{5}$$

Exogenous variables

The selection of variables aims to capture major weather and economic influences on the crop yield variability. The variables are selected according to their plant physiological impact. Across most of the subunits, but not necessarily in all, the variables are expected to be significant (see SI S.2).

Weather variables

The temporal variable division is based on an aggregated view of the plant growth process. Chmielewski and Köhn (2000) distinguish between five phenological development periods in their yield component analysis for winter rye. Butler and Huybers (2015) use for their statistical yield models four phenological development periods, while Moore and Lobell (2014) and You et al. (2009) did not divide the growing period. Dixon et al. (1994) divide the growing period by calendar months and phenological phases. They show that the division by calendar months leads to similar results in comparison to phenological phases. We distinguish two phases: the vegetative and reproductive development. The daily values of the weather variables are separately summed by calendar month over the vegetative and reproductive sections of the winter wheat and silage maize growing period. The vegetative development for winter wheat approximately lasts from November (of the planting year) to April (of the harvest year) [hereafter Nov–Apr]. For silage maize, the vegetative development has an approximate duration from May to July [May–Jul]. The reproductive development of winter wheat fulfills between May and July and that of silage maize between August and October [Aug–Oct] (DWD, 2015).

The selection of weather variables for the model is complicated by the problem of multicollinearity, i.e. a correlation between the exogenous variables. Multicollinearity leads to less precise estimates and large standard errors of the parameters. Independence among different growth variables and coverage of potential growth and stress factors is thought to be achieved by the variables: temperature normalized solar radiation (*SRT*) for the growth potential; precipitation (*PREC*) and potential evapotranspiration (*ETP*) for the water supply and the atmospheric water demand. The *SRT* is used instead of the solar radiation (R_S) to minimize the possibility of multicollinearity with the other variables. The daily *SRT* [J °C^{-1} cm^{-2}] is calculated by Eq. 6. To avoid division by zero, the temperature value is increased by 20 similar to the correction of the de Martonne aridity index (Oury, 1965).

$$SRT = \frac{R_S}{T_{\text{avg}} + 20}, \qquad \text{with} \tag{6}$$

R_S - daily solar radiation sum [J cm^{-2}],
T_{avg} - daily average temperature [°C].

The daily *ETP* [mm] is calculated following Haude (Eq. 7). *ETP* depends on the vapor pressure deficit and an empirical correction factor, the Haude factor f_H (Schrödter, 1985; Haude, 1955). For f_H, we use the arithmetic average of the values for wheat, maize, and grassland for each calendar month (see SI Tab. S.3). This considers respective characteristics of the modeled crops and is available for the calculations of all relevant months (we added the grassland values, because the values for wheat and maize are not available for the entire growing season). The vapor pressure deficit is calculated using the Magnus formula (Sonntag, 1990) by the maximum temperature (T_{max}) and the minimum temperature ($T_{\min}$) instead of dew point temperature. Thereby, we follow Castellvi et al. (1997) and Castellvi et al. (1996) who suggested the replacement of dew point temperature by $T_{\min}$ in the calculation of the ETP.

$$ETP = f_H \; 6.11 \left(\exp\left(\frac{17.269 \; T_{\max}}{237.3 + T_{\max}}\right) - \exp\left(\frac{17.269 \; T_{\min}}{237.3 + T_{\min}}\right)\right), \qquad \text{with} \tag{7}$$

f_H - Haude factor,
$T_{\max}$/ $T_{\min}$ - daily maximum/ minimum temperature.

Economic variables

Kaufmann and Snell (1997) argue that an omitted-variable bias may occur in the case of unconsidered yield-related (economic) variables in statistical models. Accordingly, we considered fertilizer price and acreage of the respective crops as economic proxy variables to control the economic yield impacts in the models. The mean annual fertilizer price serves as proxy for a set of input prices in the winter wheat and silage maize production. Price fluctuations within one year are averaged out. The effects of pre-contracts on fertilizer, which might lead to time lag effects, are neglected in our variable setting. The available data does not allow a quantification of those lag-effects yet.

Finally, our winter wheat models contain the variables: *ETP* (Nov–Apr), *ETP* (May–Jul), *SRT* (May–Jul), *PREC* (Nov–Apr), *PREC* (May–Jul), *fertilizer price*, and *acreage of winter wheat*. The silage maize models contain the variables: *ETP* (May–Jul), *ETP* (Aug–Oct), *SRT* (May–Jul), *PREC* (May–Jul), *PREC* (Aug–Oct), *fertilizer price*, and *acreage of silage maize*.

Model fit and robustness

The models should be able to reproduce both the mean level and the variability of the measured yield changes. Both characteristics are assessed by calculating the root-mean-square error (RMSE), the coefficient of determination (R^2), and the Nash–Sutcliffe efficiency coefficient (NSE) (see SI S.3 for the calculation). The RMSE captures the deviation from the mean level. The R^2 measure the reproduction of the variability. The NSE is a combined indicator for the mean model bias and the variability. It is particularly suitable for out-of-sample cross validations (Chipanshi et al., 2015; Krause et al., 2005).

The robustness of the three suggested yield models (STSM, PDM, and RCM) is assessed by running an out-of-sample cross-validation. For each year *t*, the dataset is subsequently reduced by all values of a year *t*. For this reduced dataset, we estimate the model parameters and use these parameters to calculate the yield changes of the removed years. To more rigorously check the robustness of our approach, we expanded the validation process by removing the values of further four randomly selected years, in addition to the year *t*. We refer the one-year validation simply as *validation*, while the five-year validation is called *expanded validation* hereafter.

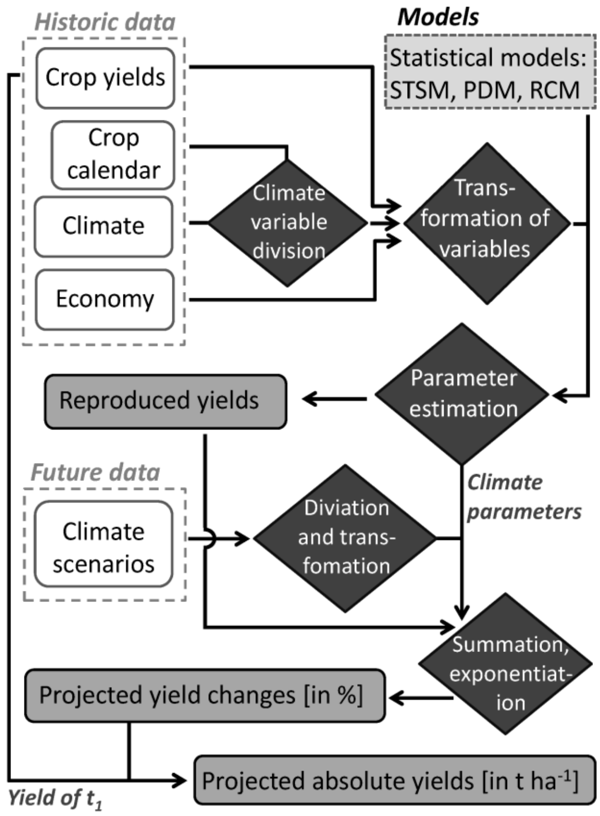
The permissibility of the OLS estimator for STSMs and PDMs is statistically tested using several tests described by Croissant and Millo (2008) and Wooldridge (2013). The regression equation specification error test (RESET) allows an evaluation whether quadratic variables are missed in the models. The Lagrange multiplier test according to Breusch–Pagan (LM) is used to examine the spatial independency (heterogeneity) of the data and justifies a spatial regression approach. Otherwise, averaging across all counties would be sufficient. The Breusch–Godfrey test is applied to assess autocorrelation and the Breusch–Pagan test is used to test for heteroskedasticity. The normal distribution of residuals is tested using the Shapiro–Wilk test. For the RCMs, there exist several criteria for the evaluation of the model goodness of fit (Reidsma et al., 2007). We use the relative criterion NSE, because it is suitable for both OLS and REML (non OLS) conditions (Reidsma et al., 2007; Chipanshi et al., 2015). A description of the applied software is given in the SI S.4.

Model application for yield projection

The calculation of the relative yield changes between the reference period (utilized for the estimation of the years t) and a projection period of the years t^* (with $t^* = M + 1, \ldots, P$) does not need an explicit specification of the basic yield level y_{t_1}. Nevertheless, it should be kept in mind that any change is related to the agronomic level of that reference. Absolute crop yields for a last year P of the projection period are calculated combining the basic yield level y_1, the yield changes during the reference period, and the projected yield changes of P following Eq. 8:

$$y_P = y_{t_1} \exp\left(\sum_{t=1}^{M} \log y'_t + \sum_{t^*=M+1}^{P} \log y'_{t^*}\right) \tag{8}$$

The concrete application is out of the scope of this paper. Here, we focus on the validation of the modeling approach. In Fig. 1a and 1b we present a schematic and algorithmic description that can be followed during applications nevertheless.



1) Estimation of the weather (β_j) and economic parameters (β_k) by Eq. 2, 3, or 4.
2) Transformation of future climate records to climate changes: $\log\, x'_{jt^*}$.
3) Multiplication and summation of the estimated β_j and the projected climate changes: $y'_{t^*} = \sum_{j=1}^{J} \beta_j \log\, x'_{jt^*}$.

Note: The intercept β_0 and the economic variables are not utilized for the projection.

4) Summation and exponentiation of the past and projected yield changes: $y'_P = \exp(\sum_{t=1}^{M} y'_t + \sum_{t^*=M+1}^{P} y'_{t^*})$.

Note: The calculation until step 4 results in a change factor that can be also converted into percentage yield changes (e.g. 1.10 → +10%).

5) Multiplication of the yield changes with

Fig. 1: (a) Workflow when using our modeling framework for projecting crop yield impacts due to climate changes and (b) steps of model building and application for accessing weather and climate impacts.

Results

Goodness of fit

STSMs, PDMs, and RCMs are able to reproduce the measured winter wheat and silage maize yield changes at the aggregated scale for the estimations and the validations (examples shown for STSMs, RCMs, and PDMs in Fig. 2). For both crops and all models, a decrease in NSE by approximately 0.25 is common when comparing estimations with validations. The NSE decreases approximately 0.38 in the expanded validation (not shown).

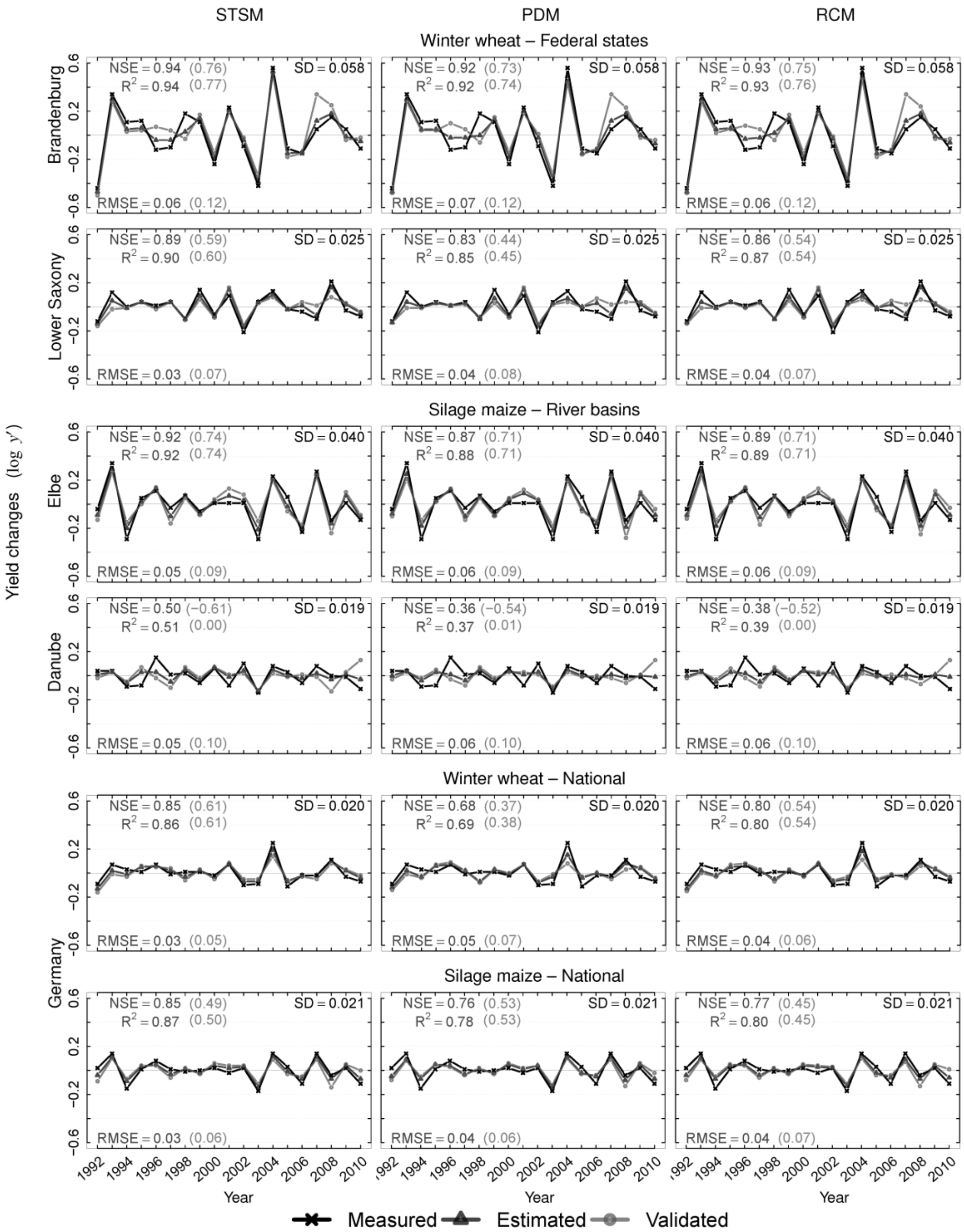


Fig. 2: Time series of measured, estimated, and validated crop yield changes of winter wheat for two German federal states, of silage maize for two river basin and for winter wheat and silage maize, respectively, at the national level using STSMs (left) , PDMs (center), and RCMs (right). The black values for NSE, R^2, and RMSE relate to the model estimation. The gray values in parenthesis characterize the model performance during validation. The SD values are the standard deviation of the measured yields.

Tab 1: The NSE measure for the winter wheat and silage maize crop yield models, the different model types (STSM, PDM, RCM), and aggregation scales (federal states, river basins, national). The arithmetic averages of federal states (FS_{Avg}) and river basins (RB_{Avg}) are given below the aggregation scales. The acronyms are: SH: Schleswig-Holstein, LS: Lower Saxony, NRW: North Rhine-Westphalia, HE: Hesse, RP: Rhineland-Palatinate, BW: Baden-Württemberg, BA: Bavaria, SL: Saarland, BB: Brandenburg, MWP: Mecklenburg-Western Pomerania, SN: Saxony, SA: Saxony-Anhalt, TH: Thuringia, DAN: Danube, ST: Schlei/ Trave, WP: Warnow/ Peene, and GER: Germany.

	Winter wheat			Silage maize		
Unit	STSM	PDM	RCM	STSM	PDM	RCM
Federal states (FS)						
SH	0.75	0.67	0.70	0.60	0.52	0.52
LS	0.89	0.83	0.86	0.78	0.65	0.71
NRW	0.84	0.80	0.81	0.65	0.53	0.57
HE	0.62	0.58	0.58	0.67	0.59	0.64
RP	0.70	0.64	0.66	0.61	0.59	0.58
BW	0.76	0.65	0.69	0.66	0.54	0.59
BA	0.72	0.55	0.64	0.67	0.51	0.58
SL	0.69	0.67	0.68	0.77	0.76	0.77
BB	0.94	0.92	0.93	0.94	0.93	0.93
MWP	0.95	0.91	0.94	0.93	0.90	0.92
SN	0.85	0.73	0.79	0.72	0.58	0.67
SA	0.89	0.86	0.89	0.93	0.92	0.92
TH	0.67	0.63	0.65	0.81	0.78	0.79
FSAvg	0.79	0.73	0.76	0.75	0.68	0.71
River basins (RB)						
Eider	0.75	0.72	0.75	0.43	0.33	0.37
ST	0.69	0.64	0.67	0.54	0.50	0.52
Elbe	0.91	0.81	0.88	0.92	0.87	0.89
Weser	0.88	0.82	0.85	0.84	0.78	0.77
Ems	0.85	0.78	0.83	0.57	0.34	0.50
Rhine	0.78	0.68	0.73	0.76	0.67	0.69
Maas	0.69	0.68	0.68	0.75	0.74	0.75
DAN	0.72	0.62	0.66	0.50	0.36	0.38
WP	0.94	0.91	0.94	0.93	0.90	0.92
Oder	0.75	0.77	0.77	0.93	0.91	0.92
RSAvg	0.80	0.74	0.78	0.72	0.64	0.67
National						
GER	0.85	0.68	0.80	0.85	0.76	0.77

For all models, crops, and aggregation scales, the goodness of fit (measured by the NSE for the aggregated scales) is shown in Tab. 1. Generally, the NSE decreases from STSM over RCM to PDM for both winter wheat and silage maize in the estimations. For both crops, there exists a significant Pearson correlation ($^{***}p \leq 0.01$, $^{**}p \leq 0.05$, $^{*}p \leq 0.1$, $p > 0.1$) between yield variability (i.e., standard deviation (SD)) and goodness of fit (i.e., NSE). However, this correlation is stronger for silage maize (0.86^{***}) than for winter wheat (0.66^{**}) on the federal state scale. This difference is also illustrated by the fact that the SDs of both crops are strongly correlated (0.93***), but the corresponding NSEs correlate only by 0.64**.

Differences between models exist in their reproduction of extreme inter-annual yield changes in single counties (Fig. 3). In the federal state of Brandenburg, for example, the STSMs (NSE: 0.84) reproduce extreme county yield changes (see county LOS: Oder-Spree) considerably better (light yellow) than RCMs (NSE: 0.81) and PDMs (NSE: 0.70) (indicated by the blue, red, and orange color).

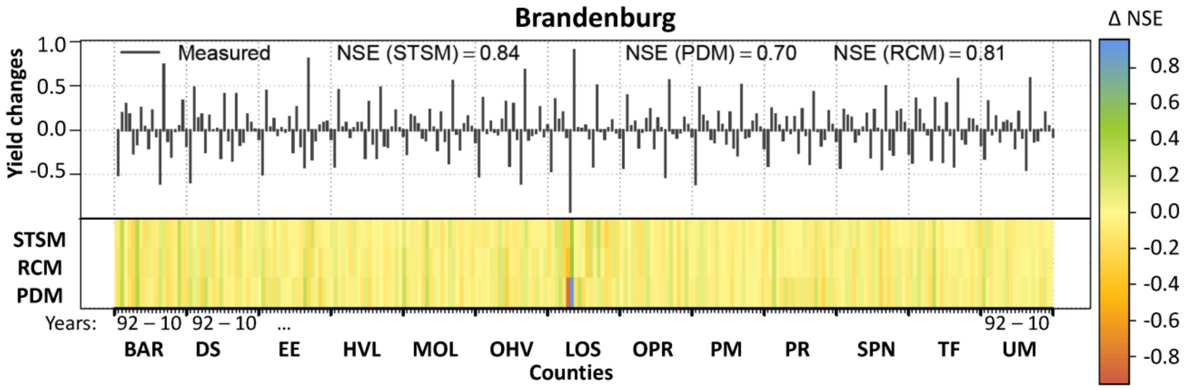


Fig. 3: Time series of measured winter wheat yield changes (bars) and the difference ($\Delta \log y' = \log y'_{measured} - \log y'_{estimated}$) between measured and estimated winter wheat yield changes for STSM, PDM, and RCM (bottom). The data records for each county last from 1992 to 2010. The county acronyms are: Barnim (BAR), Dahme-Spreewald (DS), Elbe-Elster (EE), Havelland (HVW), Märkisch-Oderland (MOL), Oberhavel (OHV), Oder-Spree (LOS), Ostprignitz-Ruppin (OPR), Potsdam-Mittelmark (PM), Prignitz (PR), Spree-Neiße (SPN), Teltow-Fläming (TF), and Uckermark (UM).

Aggregation effect

The aggregation of estimated yield changes generally increases the goodness of fit from the county to the federal state, river basin, and national scale. The extent of this improvement is shown in Tab. 2. PDMs, the models with the lowest goodness of fit measured with the NSE at the county level, gained the highest accuracy improvement by aggregation. This aggregation effect is similar for RCMs and STSMs. The aggregation to the national scale has the highest average (for all models) aggregation effect (+0.29), for river basins and federal states it is nearly the same (+0.16).

Tab. 2: Effect of spatial aggregation expressed as difference in the NSE values (ΔNSE) between the NSE at the aggregated scale (units) and the mean of *N* basic NSE_i across *N* counties: $\Delta\text{NSE} = \text{NSE}_{Unit} - N^{-1}\sum_{i=1}^{N}\text{NSE}_i$. Other terms similar to Tab. 1.

	Winter wheat			Silage maize		
Unit	STSM	PDM	RCM	STSM	PDM	RCM
Federal states (FS)						
SH	0.08	0.15	0.09	0.11	0.23	0.14
LS	0.13	0.26	0.16	0.15	0.30	0.15
NRW	0.16	0.36	0.25	0.15	0.39	0.22
HE	0.05	0.21	0.12	0.23	0.52	0.20
RP	0.08	0.22	0.10	0.11	0.32	0.17
BW	0.10	0.16	0.10	0.11	0.27	0.19
BA	0.14	0.33	0.15	0.14	0.22	0.12
SL	0.06	0.12	0.08	0.05	0.24	0.03
BB	0.09	0.22	0.12	0.08	0.14	0.08
MWP	0.13	0.29	0.17	0.12	0.22	0.14
SN	0.09	0.17	0.05	0.12	0.16	0.07
SA	0.04	0.20	0.06	0.11	0.19	0.09
TH	0.11	0.27	0.20	0.15	0.25	0.17
FSAvg	0.10	0.23	0.13	0.13	0.27	0.14
River basins (RB)						
Eider	0.07	0.16	0.13	0.04	0.12	0.04
ST	0.02	0.07	0.02	0.07	0.15	0.12
Elbe	0.18	0.43	0.17	0.20	0.38	0.14
Weser	0.16	0.29	0.20	0.21	0.45	0.22
Ems	0.12	0.23	0.16	0.05	0.20	0.03
Rhine	0.14	0.36	0.19	0.20	0.49	0.23
Maas	0.02	0.15	0.05	0.18	0.33	0.21
DAN	0.18	0.44	0.18	0.04	0.18	0.08
WP	0.11	0.28	0.12	0.12	0.23	0.11
Oder	0.01	0.06	0.05	0.02	0.06	0.02
RBAvg	0.10	0.25	0.13	0.11	0.26	0.12
National						
GER	0.19	0.44	0.21	0.27	0.49	0.16

An example of the consequences of the aggregation for the goodness of fit distributions is depicted for the winter wheat STSMs in Fig. 4. Aggregation, on the one hand, leads to a coarser resolution; on the other hand, it improves the goodness of fit for larger regions. The aggregation from county to the river basin scale shows a west-east gradient from lower to higher NSEs. The aggregation to federal states shows that the Central Uplands between Saarland and Thuringia achieve the lowest NSE values. The statistical significant STSMs are shown Fig. 4. The PDMs are all significant at $p \leq 0.05$ (*F*-test).

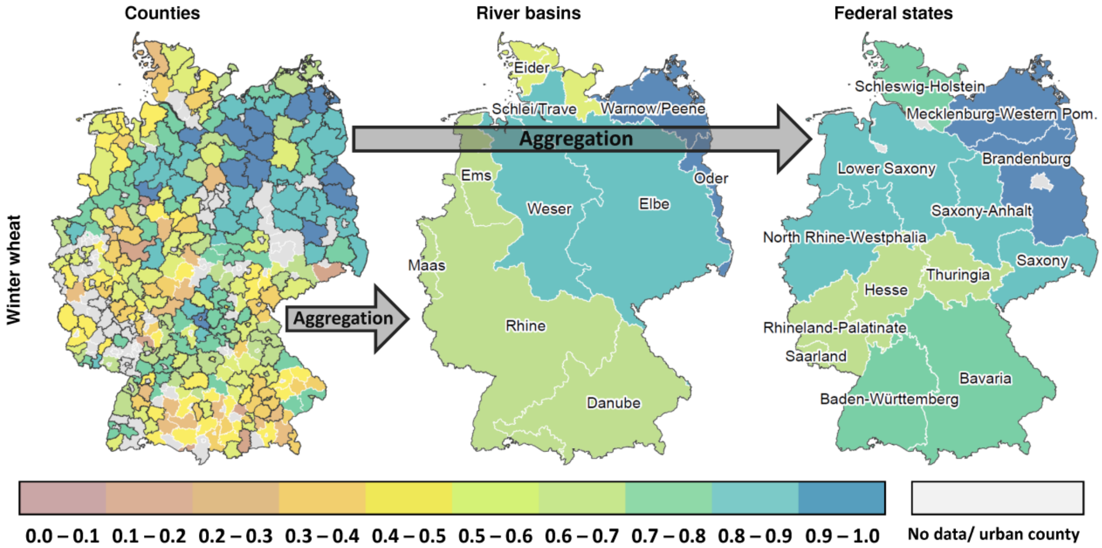


Fig. 4: Spatial distribution of the NSE for STSMs estimated for winter wheat at different scales (counties, *N*=289; federal states, *N*=13; river basins, *N*=10). Counties with significant effects (*F*-test, $p \leq 0.10$) STSMs (*N*=189) are bordered black (left).

Parameter heterogeneity of weather variables

The parameter distribution of all models, across scales, and for both crops is shown in Fig. 5. Due to the chosen functional form, the parameters are directly comparable. A wide range of the boxplots reflects a high spatial heterogeneity. The ranges of the parameter values are generally substantially smaller for the PDMs and the RCMs than for the STSMs. Some variables clearly show diversions from zero (*ETP* May–Jul and *SRT* May–Jul for winter wheat and *ETP* May–Jul for silage maize) while others are distributed around zero.

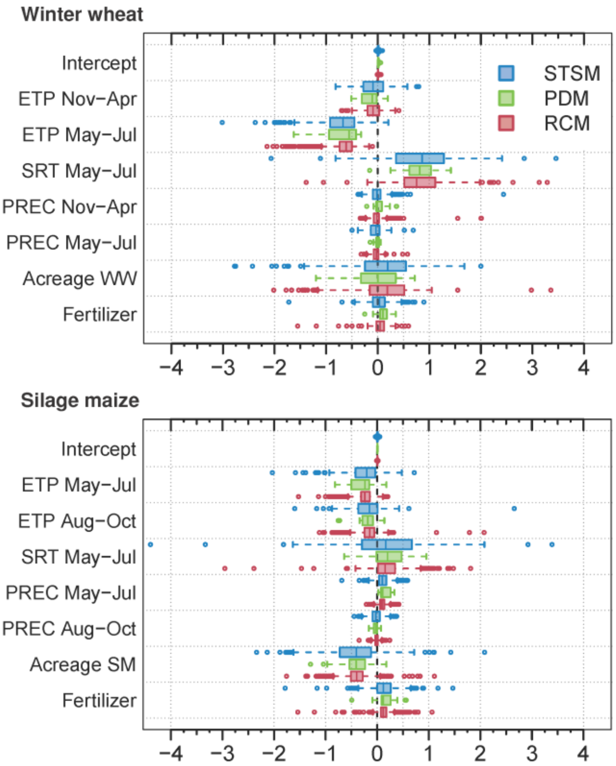


Fig. 5: Parameter distributions across counties, river basins and federal states of separately estimated time series models (STSMs), panel data models (PDMs), and random coefficient model (RCMs). The models are applied to single counties (STSMs), river basins (PDMs, RCMs) and federal states (PDMs, RCMs) for winter wheat and silage maize. The band inside the box is the median, the box represents the 25% and 75% quartile. The whiskers are defined as the maxima and the minima as long as both values are within the 1.5 interquartile-range from the median. Otherwise this range is shown and outliers outside the range are depicted as points.

For the STSMs, the spatial heterogeneous parameter variation is depicted in Fig. 6a-t. The maps of the county-individual (Fig. 6, left two columns) and the per-federal-state-averaged parameter values (Fig. 6, right two columns) often show spatial pattern, which scatter around the parameter main tendencies in Fig. 5. The larger patterns are easier to reveal following the maps with the averaged values. In particular, several parameter maps contain east-west patterns with stronger effects in eastern than in the western federal stats. That is the case for the winter wheat variables *PREC* Nov–Apr (Fig. 6 m, o), *ETP* Nov–Apr (Fig. 6 a, c), and *ETP* May–Jul (Fig. 6 e, g), and to a lesser extent for the silage maize variables *PREC* May–Jul (Fig. 6 n, p), *ETP* May-Jul (Fig. 6 b, d), and *ETP* Aug–Oct (Fig. 6 f, h). For winter wheat and silage maize, the patterns of the *SRT* May–Jul parameter reveal a north-south gradient with stronger effects in the north and weaker in the south, whereby this is more expressed for wheat than for maize.

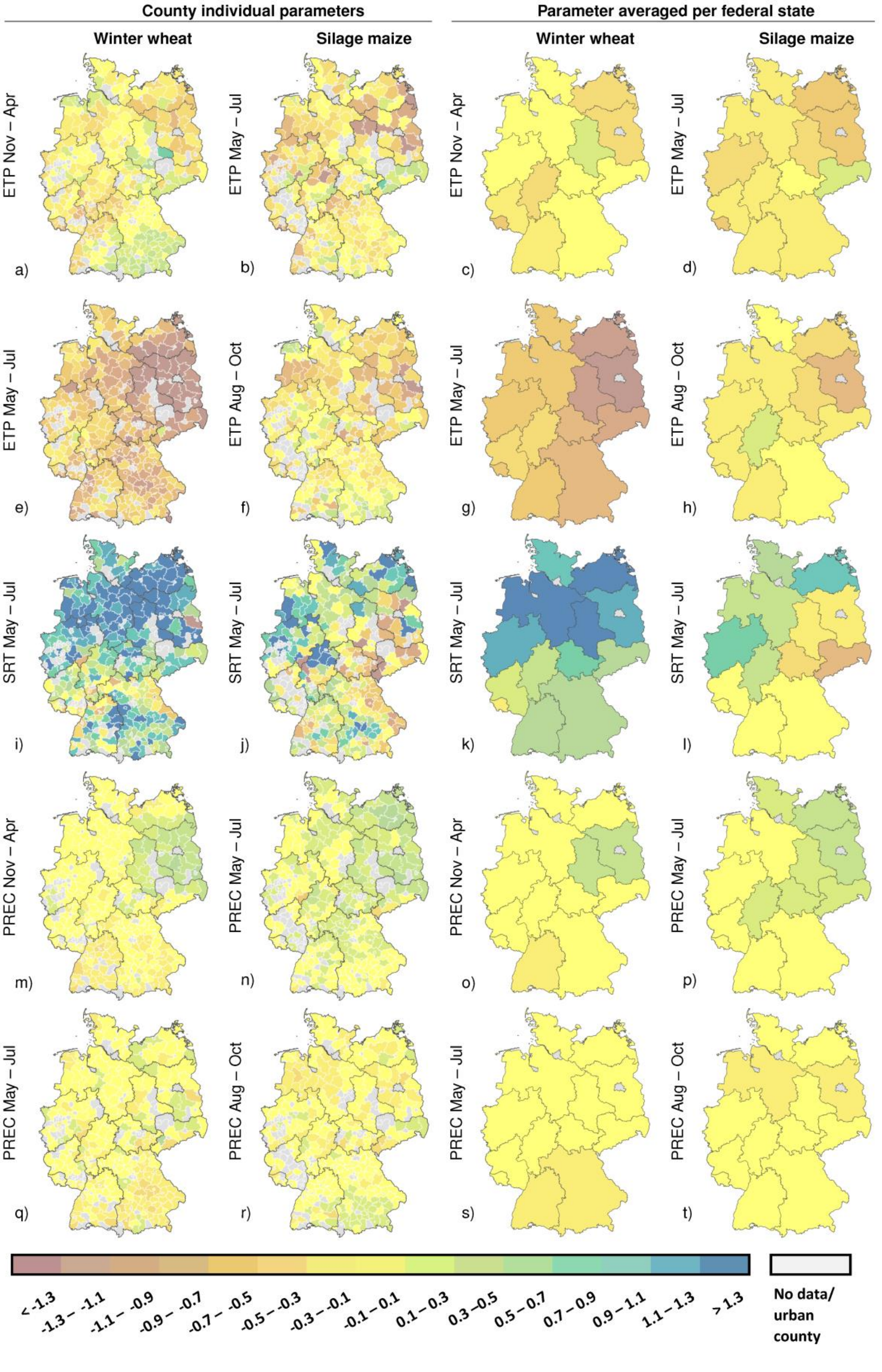


Fig. 6: Spatial distribution of the STSM parameter values for winter wheat and silage maize across counties (left two columns) and averaged per federal state (right two columns). The aggregated federal states parameter values are estimated on county scale (Eq. 2) and aggregated by arithmetic average in hindsight. The weather parameters of winter wheat are ETP Nov–Apr (a,c), ETP May–Jul (e,g), SRT May–Jul (i,k), PREC Nov–Apr (m,o), and PREC May–Jul (q,s). The weather parameters of silage maize are ETP May–Jul (b,d), ETP Aug–Oct (f,h), SRT May–Jul (j,l), PREC May–Jul (n,p), and PREC Aug–Oct (r,t).

Statistical tests

The validity of a spatially distributed modeling approach is statistically confirmed by the LM-test. The county yields do not dependent on each other. The Cobb–Douglas function is mostly not error-specified as the functional form (RESET) when testing the STSMs of both crops (SI Fig. S.1). This means quadratic variables would not improve the model goodness of fit. According to the RESET, the winter wheat PDMs are partly error-specified and the silage maize PDMs are often error-specified (SI

Tab. S.1). Both autocorrelation and heteroskedasticity do not occur in the majority of the STSMs. Their residuals are mostly normally distributed (SI Fig. S.1). For the PDMs, autocorrelation and/ or heteroskedasticity are common (SI Tab. S.1). The quality of the RCMs can be assessed using the NSEs in Tab. 1.

Several highly significant correlations exist among the transformed, weather and economic variables. For instance, *SRT* is moderately correlated with *ETP* and precipitation (0.67^{***}, -0.52^{***}), as well as precipitation and *ETP* (-0.58^{***}), in the period May to Jul. Furthermore, fertilizer price and acreage of winter wheat and silage maize are strongly correlated (0.72^{***}, 0.87^{***}). However, a test of multicollinearity (condition index) resulted in values always lower than 12 (lower 8 in 97.5% of all cases). Following Belsley et al. (1980), values beyond 30 are an indicator for multicollinearity. Further information about the results of the statistical tests and the correlation of all exogenous variables are presented in the SI Fig. S.1, Tab. S.1 and S.2.

Discussion

Goodness of fit and yield variability between crops and regions

We investigate and validate three statistical models (STSM, PDM, and RCM) according their robustness for short and medium-term yield assessments. These three models are tested to capture the weather -related yield variability of winter wheat and silage maize in Germany. All models are able to satisfactorily reproduce the temporal and spatial variability of yields. In general, the differences in goodness of fit between winter wheat and silage maize are low. The models of regions with higher yield variability have generally scored the highest NSE. Thus, a clear west–east NSE gradient for the federal state and river basin scale is observable. We found a very strong positive correlation between the goodness of fit and yield variability for silage maize (0.86) and a strong positive correlation for winter wheat (0.66). This relationship is visible across the federal states, but not between the two crops.

The STSMs, which are the simplest models, perform best, followed by RCMs and PDMs. The ranking holds true for estimations, validations, and expanded validations, which indicates that STSMs are robust to missing data despite their higher parameter numbers. The advantage in robustness of STSMs compared to the other two approaches originates in the estimation method. The STSMs are separately estimated for each county. The parameters of STSMs vary more across the counties than those of RCMs and PDMs. Thus, STSMs can reproduce better extreme county individual yield anomalies than RCMs and PDMs, which might be beneficial for the adequate reproduction of heterogeneous spatial conditions (Beck and Katz, 2007; Butler and Huybers, 2013). In contrast, the PDM parameters are estimated for the entire dataset to reproduce the full range of yield variability. Their parameters reproduce rather the mean yield level than the full range yields variability. Nevertheless, the higher parameter number and the lower degree of freedom of the STSMs are a potential source of parameter instability when the time series becomes shorter. Under such conditions, RCMs and PDMs might behave more robustly than STSMs (Reidsma et al., 2007; You et al., 2009). Furthermore, projections at the county scale beyond the observed yield variability might be more biased by single regional events at that scale when using STSMs compared to PDMs and RCMs. Such a disadvantage might lose relevance with increasing spatial aggregation of the STSM results.

Aggregation effect

A substantially higher goodness of fit is achieved by all models after aggregation from county to larger spatial units. The winter wheat STSMs (+0.10) show the smallest aggregation effect in comparison to the respective RCMs (+0.13) and to the PDMs (+0.23) at the averaged federal state scale. For silage maize and the other scales, the effects are similar. In tendency, the advantages of STSMs at the lowermost scale (here county) lose relevance at the more aggregated scale. Thus, aggregation has a slighter effect on the goodness of fit for STSMs than of RCMs and PDMs. Woodard and Garcia (2008), Lobell and Burke (2010), and Hanus (1978) have noted the aggregation effect before. Conradt et al. (2015) used the parameter vectors of our STSMs for cluster analyses to define optimized PDM aggregations independently from federal states or river basin scales. This could again, but only slightly, add to the overall goodness of fit; at least the county-specific fidelity of the estimations became much more homogeneous. In our approach, only the estimated outcomes are spatially aggregated and not the exogenous variables. Aggregated exogenous variables can lead to an underestimation of the weather effect (Garcia et al., 1987), to decreased variability, and erroneous results (Finger, 2012).

Parameter distributions and patterns

Winter wheat is more responsive than silage maize to higher evaporative demand during spring and summer as indicated by the more negative values for *ETP* May–Jul. That might be due to the more developed plant canopy. After closing the canopy (Aug–Oct), the silage maize shows a clearer negative impact of higher *ETP*. For the *ETP* related vapor pressure deficit, a negative yield impact is also shown by Lobell et al. (2014) and Roberts et al. (2012). Consistent with this explanation, the less developed silage maize in May to Jul (early vegetative development, between emergence and canopy closure) is more sensitive to lower water supply than winter wheat during that time (*PREC* May–Jul). For winter wheat, a similar effect is observable during the early plant development stages, in particular in the eastern parts of Germany. This region is marked by sandy soils with low water holding capacity and low precipitation levels. These conditions lead to a higher sensitivity of crop yields (high yield variability) to inter-annual changes of water supply. Wessolek and Asseng (2006) also show the importance of this limited water supply for winter wheat in north-east Germany. The importance of the water supply for winter wheat and silage maize in Germany is also emphasized by Kersebaum and Nendel (2014) and Wolf and Diepen (1994).

Furthermore, winter wheat benefits more than silage maize from higher *SRT* May–Jul values during that period. This might reflect the higher temperature sensitivity of light respiration of C_3- (e.g. wheat) than C_4-crops (e.g. maize). As a consequence, lower temperatures at the same radiation levels and higher radiation levels at the same temperature levels (increasing *SRT*) function more positively on winter wheat than on silage maize (Long et al., 2006; Rötter and van de Geijn, 1999). However, Conradt et al. 2015 could increase the model performance by decoupling radiation and temperature back into two model parameters considering regional exceptions to these general patterns. A detailed discussion of the spatial parameter patterns for winter wheat and silage maize is in the SI S.6.

Surprisingly, the parameters for precipitation indicate a small yield effect compared to the other factors considered for both crops. A possible explanation is offered by the variability differences among variables. In our dataset, the transformed precipitation from May to Jul varies by ±43% (relative SD), while *ETP* and *SRT* only vary by ±17% and ±8%, respectively. Due to their high relative SD, small parameter values are estimated for precipitation. However, for the assessment of weather-yield impacts the explained yield variability ($\beta_j \ \log x_j'$) is more important than the parameter size (β_j). A further analysis shows that the yield variability explained by precipitation is substantially larger in comparison to the other variables (SI Fig. S.2). The result possibly explains the small yield impact for precipitation in Europe reported by Moore and Lobell (2014). They have drawn their conclusion solely from the parameter size, but not from the explained variability.

Generally, the STSM parameters show parameter patterns on a broader scale, but also county specific heterogeneity. The spatial parameter patterns can be explained by linear relationships between yield and exogenous variables, because of spatially heterogeneous levels of the exogenous variables. The county parameters do not deviate ideally from the average parameters of the broader patterns. In our case, the parameters may also reflect individual factor influences, which are not considered in the model. These influences are collinear with the considered variables but not relevant in the majority of the counties (county individual time variant effects). The impact of those factors may lead to spatial heterogeneity between neighboring counties that cannot be explained by differences in soil characteristics or cropping structure. For instance, the possible collinear influences might be catch crops (*ETP*), weeds, pests, and diseases (*SRT*), or irrigation (*PREC*).

Model application in climate impact studies

Our modeling scheme allows a direct interpretation of the spatial parameter variability and a usage for crop yield assessments with seasonal- and medium-term climate projections. Both characteristics are based on a consequent usage of changes instead of absolute values, which contributes to the methodological novelty of the approach. The parameter values and patterns can be used to prove the plausibility of the model outcome. The feasibility of plausibility test is supported by a variable definition that reflects major climate impacts on potential growth and stress related to limited water supply. The selected variables might be meaningful also in other wheat and maize growing regions. However, an adjustment of the temporal division to the regional crop calendar is necessary. The use of changes makes the model also insensitive to systematic errors in data from climate simulations. This insensitivity does not avoid flawed yield projections of flawed climate simulations. However, considering the necessary effort of bias correction and the often nontransparent procedure (Lobell, 2013), our models are an option for using the outcome of climate simulations in advance of a later bias correction. It is not the solution for the bias problem of climate simulations, but an improvement for their technical handling.

Butler and Huybers (2013) show that the impact of temperature on US maize yields is very sensitive in respect to the latitude and the regional climate conditions. STSMs, PDMs and RCMs should be principally applicable for such conditions in order to project climate impacts. However, the advantages and limitations of each model should be kept in mind. Our approach implicitly accounts for the different yield sensitivities of vegetative and reproductive growth periods to climate changes. Any further detailed resolution of the phenological development might be beneficial in statistical analysis (Butler and Huybers, 2015). However, yield projection of statistical yield models would require phenological development data also for the future. Since phenological models (Ma et al., 2012) and climate simulations (Lobell, 2013) are becoming robust only at broader temporal and spatial resolution, we use monthly averaged phenological dates, to make our models suitable for future projections.

Our statistical models project future yield changes on the basis of the current system. Several factors and factor relationships that are unknown today might play a major role in the future and are not included in the model. In our model set-up, we focused on the representation of regularly returning yield impacts of climate variables that can be reliably received from climate models. The impact of extreme weather events that affected the crop yield only episodically in the past but will become regular disturbances in the future might be underestimated. Furthermore, if climatic change passes thresholds, crop yields might be seemingly insensitive due to unconsidered climate impacts during the parameter estimation of our crop yield models (Blanc and Sultan, 2015; Rötter et al., 2011). Yield effects of technological change and the impact of higher CO_2 (by stimulating crop growth and increasing water use efficiency) are also not included in the model. They could be taken into account by introducing a post-processing to the model output using external correction factors as exercised by Wechsung et al. (2008).

Conclusion

Our suggested approach can be used for seasonal yield forecasts and climate impact projections on crop yields. For short and medium term climate assessments, we investigate and validate three types of statistical crop yield models (STSM, PDM, and RCM). These models are suitable for a combination with biased climate simulations and avoid explicit modeling of crop yield trends. Our approach is thoroughly based on relative changes of yields and yield influencing factors. Our models can reproduce past regional yield variability; they are robust to data fragmentation and show reasonable parameter patterns at aggregated scales. Although STSMs have shown the best performance at the aggregated scale, the model assessments at the county scale should only be used as technical intermediate steps but not as projections. The suggested regression models might be applicable to calculate weather-related yield risks and thus support investment decisions (e.g. in irrigation systems) and risk pricing (e.g. of harvests or agricultural commodity futures for farmers, traders, or insurance companies).

References

Angulo, C. et al., 2013. Implication of crop model calibration strategies for assessing regional impacts of climate change in Europe. Agricultural and Forest Meteorology, 170: 32-46.

Asseng, S. et al., 2013. Uncertainty in simulating wheat yields under climate change. Nature Climate Change, 3: 827–832.

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

Beck, N. and Katz, J.N., 2007. Random Coefficient Models for Time-Series–Cross-Section Data: Monte Carlo Experiments. Political Analysis, 15: 182-195.

Belsley, D.A., Kuh, E. and Welsch, R.E., 1980. Regression diagnostics: identifying influential data and sources of collinearity. Wiley, New York, 300 p.

Blanc, E. and Sultan, B., 2015. Emulating maize yields from global gridded crop models using statistical estimates. Agricultural and Forest Meteorology, 214–215: 134-147.

Butler, E.E. and Huybers, P., 2013. Adaptation of US maize to temperature variations. Nature Climate Change, 3: 68-72.

Butler, E.E. and Huybers, P., 2015. Variations in the sensitivity of US maize yield to extreme temperatures by region and growth phase. Environmental Research Letters, 10: 1-8.

Castellvi, F., Perez, P.J., Stockle, C.O. and Ibaiiez, M., 1997. Methods for estimating vapor pressure deficit at a regional scale depending on data availability. Agricultural and Forest Meteorology, 87(4): 243-252.

Castellvi, F., Perez, P.J., Villar, J.M. and Rosell, J.I., 1996. Analysis of methods for estimating vapor pressure deficits and relative humidity. Agricultural and Forest Meteorology, 82: 29-45.

Chipanshi, A. et al., 2015. Evaluation of the Integrated Canadian Crop Yield Forecaster (ICCYF) model for in-season prediction of crop yield across the Canadian agricultural landscape. Agricultural and Forest Meteorology, 206: 137-150.

Chmielewski, F. and Köhn, W., 2000. Impact of weather on yield components of winter rye over 30 years. Agricultural and Forest Meteorology, 102: 253-261.

Conradt, T., Gornott, C. and Wechsung, F., 2015. Extending and improving regionalized winter wheat and silage maize yield regression models for Germany: enhancing the predictive skill by panel definition through cluster analysis. Agricultural and Forest Meteorology, (216) 68–81.

Croissant, Y. and Millo, G., 2008. Panel Data Econometrics in R: The plm Package. Journal of Statistical Software, 27(2): 1-43.

Dielman, T.E., 1983. Pooled Cross-Sectional and Time Series Data: A Survey of Current Statistical Methodology. The American Statistician, 37(2): 111-122.

Dixon, B.L., Hollinger, S.E., Garcia, P. and Tirupattur, V., 1994. Estimating Corn Yield Response Models to Predict Impacts of Climate Change. Journal of Agricultural and Resource Economics, 19(1): 58-68.

DWD, 2011. Daily weather data, 1951 - 2010. German Weather Service.

DWD, 2015. Aktueller Stand der Phänologie in Deutschland, 1992 - 2013. German Weather Service, Retrieved from: http://www.dwd.de/phaenologie (11.06.2015).

Estes, L.D. et al., 2013. Projected climate impacts to South African maize and wheat production in 2055: a comparison of empirical and mechanistic modeling approaches. Global Change Biology, 19(12): 3762–3774.

Federal Statistical Office, 2013. Ackerland nach Hauptfruchtgruppen und Fruchtarten.

Finger, R., 2012. Biases in Farm-Level Yield Risk Analysis due to Data Aggregation. German Journal of Agricultural Economics, 61: 30-41.

Garcia, P., Offutt, S.E., Pinar, M. and Changnon, S.A., 1987. Corn Yield Behavior - Effects of Technological Advance and Weather-Conditions. Journal of climate an applied meteorology, 26: 1092-1102.

Gornott, C. and Wechsung, F., 2015. Niveauneutrale Modellierung der Ertragsvolatilität von Winterweizen und Silomais auf mehreren räumlichen Ebenen in Deutschland. Journal für Kulturpflanzen, 67: 205-223.

Hanus, H., 1978. Vorhersage von Ernteerträgen aus Witterungsdaten in den Ländern der EG. eurostat, 21: i-54.

Haude, W., 1955. Zur Bestimmung der Verdunstung auf möglichst einfache Weise. Mitteilungen des Deutschen Wetterdienstes 11.

Iizumi, T. et al., 2013. Prediction of seasonal climate-induced variations in global food production. Nature Climate Change, 3: 904–908.

Kaufmann, R.K. and Snell, S.E., 1997. A Biophysical Model of Corn Yield: Integrating Climatic and Social Determinants. American Journal of Agricultural Economics, 79(1): 178-190.

Kersebaum, K.C. and Nendel, C., 2014. Site-specific impacts of climate change on wheat production across regions of Germany using different CO2 response functions. European Journal of Agronomy, 52: 22-32.

Krause, J., 2008. A Bayesian approach to German agricultural yield expectations. Agricultural Finance Review, 68: 9-23.

Krause, P., Boyle, D.P. and Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. Advances in Geosciences, 5: 89-97.

Lee, B.-H., Kenkel, P. and Brorsen, B.W., 2013. Pre-harvest forecasting of county wheat yield and wheat quality using weather information. Agricultural and Forest Meteorology, 168: 26-35.
Lobell, D.B., 2007. Changes in diurnal temperature range and national cereal yields. Agricultural and Forest Meteorology, 145: 229–238.
Lobell, D.B., 2013. Errors in climate datasets and their effects on statistical crop models. Agricultural and Forest Meteorology, 170: 58-66.
Lobell, D.B. and Asner, G.P., 2003. Climate and Management Contributions to Recent Trends in U.S. Agricultural Yields. Science, 299(5609): 1032.
Lobell, D.B. and Burke, M.B., 2010. On the use of statistical models to predict crop yield responses to climate change. Agricultural and Forest Meteorology, 150(11): 1443-1452.
Lobell, D.B. et al., 2013. The critical role of extreme heat for maize production in the United States. Nature Climate Change, 3(5): 497-501.
Lobell, D.B. et al., 2014. Greater Sensitivity to Drought Accompanies Maize Yield Increase in the U.S. Midwest. Science, 344(6183): 516-519.
Long, S.P., Zhu, X.G., Naidu, S.L. and Ort, D.R., 2006. Can improvement in photosynthesis increase crop yields? . Plant Cell and Environment, 29: 315-330.
Ma, S., Churkina, G. and Trusilova, K., 2012. Investigating the impact of climate change on crop phenological events in Europe with a phenology model. International Journal of Biometeorology, 56(4): 749-763.
Monteith, J.L., 1977. Climate and the efficiency of crop production in Britain. Philosophical Transactions of the Royal Society of London, 281: 277-294.
Moore, F.C. and Lobell, D.B., 2014. Adaptation potential of European agriculture in response to climate change. Nature Climate Change, 4: 610–614.
Mueller, N.D. et al., 2012. Closing yield gaps through nutrient and water management. Nature, 490: 254-257.
Oury, B., 1965. Allowing for Weather in Crop Production Model Building. Journal of Farm Economics, 47(2): 270-283.
Palosuo, T. et al., 2011. Simulation of winter wheat yield and its variability in different climates of Europe: A comparison of eight crop growth models. European Journal of Agronomy, 35(3): 103-114.
Ray, D.K., Gerber, J.S., MacDonald, G.K. and West, P.C., 2015. Climate variation explains a third of global crop yield variability. Nature Communications, 6: 1-9.
Reidsma, P., Ewert, F. and Lansink, A.O., 2007. Analysis of farm performance in Europe under different climatic and management conditions to improve understanding of adaptive capacity. Climatic Change, 84(3-4): 403-422.
Roberts, M.J., Schlenker, W. and Eyer, J., 2012. Agronomic weather measures in econometric models of crop yield with implications for climate change. American Journal of Agricultural Economics, 95(2): 236-243.
Rötter, R. and van de Geijn, S.C., 1999. Climate Change Effects on Plant Growth, Crop Yield and Livestock. Climatic Change, 43: 651-681.
Rötter, R.P., Carter, T.R., Olesen, J.E. and Porter, J.R., 2011. Crop-climate models need an overhaul. Nature Climate Change, 1: 175–177.
Schlenker, W. and Roberts, M.J., 2009. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. Proceedings of the National Academy of Sciences of the United States of America, 106(37): 15594-15598.
Schrödter, H., 1985. Verdunstung: Anwendungsorientierte Meßverfahren und Bestimmungsmethoden. Springer, 204 p.
Sonntag, D., 1990. Important new Values of the Physical Constants of 1986, Vapour Pressure Formulations based on ITS-90, and Psychrometer Formulae. Meteorologische Zeitschrift, 4(5): 340-344.
Statistical Offices of the Federation and the Länder, 2013a. Datensatz Anbaufläche (Feldfrüchte und Grünland): Deutschland, Jahre, Fruchtarten 1991-2007.
Statistical Offices of the Federation and the Länder, 2013b. Hektarerträge ausgewählter landwirtschaftlicher Feldfrüchte - Jahressumme - regionale Tiefe: Kreise und krfr. Städte.
Statistical Offices of the Federation and the Länder, 2013c. Einkaufspreise landwirtschaftlicher Betriebsmittel.
Wechsung, F., Lüttger, A.B. and Hattermann, F., 2008. Projektionen zur klimabedingten Änderung der Erträge von einjährigen Sommer-und Winterkulturen des Ackerlandes am Beispiel von Silomais und Winter-weizen, Die Ertragsfähigkeit ostdeutscher Ackerflächen unter Klimawandel. PIK-Report 112, p. 18-32.
Wessolek, G. and Asseng, S., 2006. Trade-off between wheat yield and drainage under current and climate change conditions in northeast Germany. European Journal of Agronomy, 24(4): 333-342.
Wolf, J. and Diepen, C.A.v., 1994. Effects of climate change on silage maize production potential in the European Community. Agricultural and Forest Meteorology, 71: 33-60.
Woodard, J.D. and Garcia, P., 2008. Weather Derivatives, Spatial Aggregation, and Systemic Risk: Implications for Reinsurance Hedging. Journal of Agricultural and Resource Economics, 33(1): 34-51.
Wooldridge, J.M., 2013. Introductory econometrics. a modern approach. South Western Cengage Learning, 868p.
You, L., Rosegrant, M.W., Wood, S. and Sun, D., 2009. Impact of growing season temperature on wheat productivity in China. Agricultural and Forest Meteorology, 149(6-7): 1009-1014.

Supplemental Information

Statistical regression models for assessing climate impacts on crop yields – A validation study for winter wheat and silage maize in Germany

Christoph Gornott[1*] and Frank Wechsung[1]

[1] Potsdam Institute for Climate Impact Research (PIK)

[*] Corresponding author

Data and aggregation

Winter wheat and silage maize yield data are available on county scale for the period from 1991 to 2010. The yields between 1991 and 1998 are digitized from the statistical yearbooks of the German federal states. The yields between 1999 and 2010 are digitally available from the Statistical Offices of the Federation and the Länder (2013b). The time series for Saxony last only from 1992 to 2007 and those for Saxony-Anhalt from 1991 to 2006. Counties without or with incomplete yield data are not considered in our analysis. The weather data contains temperature as daily maximum (T_{max}), minimum (T_{min}), and average (T_{avg}) and solar radiation (R_S) and precipitation as daily sums. The weather data are based on measurements at 1,218 weather stations of the German weather service (DWD, 2011) within Germany. The weather stations are assigned to the counties according to their location. In case of more than one weather station per county, we take the arithmetic average of all stations. Counties without weather stations and the weather stations above an altitude of 700m are unconsidered (6.9% of the 1,218 weather stations). This altitude restriction is chosen because husbandry is not practiced above this altitude in Germany. The economic proxy variables acreage and fertilizer price are observed only on national scale (for Germany). The acreage data of winter wheat and silage maize is based on datasets of the Statistical Offices of the Federation and the Länder (2013a) [1991 to 2008] and the Federal Statistical Office (2013) [2008 to 2010]. The fertilizer price (and further factor and product prices) is from the Statistical Offices of the Federation and the Länder (2013c).

Using statistically not significant variables

Statistical significance is not the only criteria for the variable selection (Wooldridge, 2013, p. 127-129). Nuzzo (2014) shows that results can be made also more plausible by a low p-value. A high statistical significance means that the probability of the correct result increases. Wooldridge (2013, p. 141) describes that individually statistically significant variables in combination with other variables are often no longer significant and *vice versa*. Studenmund (2000, p. 172-173) criticizes that a stepwise regression, which takes successively significant variables in a model, is “an admission of ignorance" of the variable selection. The arbitrary order to select variables prevents a plant-physiologically reasonable selection of the variables.

Prost et al. (2008) and Whittingham et al. (2006) show the limitations of a stepwise regression, because of the variable section can be biased due to the selection procedure and the selection criteria. Furthermore, the variable selection highly depends on the estimation dataset and is only limited exportable to other datasets, regions ore time periods. Finally, important variables, like the precipitation, are occasionally not considered by the stepwise approach. In such a case, the projections might be affected by an omitted variable bias.

Model fit

The NSE (Eq. S.1), the R^2 (Eq. S.2), and the RMSE (Eq. S.3) are calculated by the estimated (E) and observed (O) yield changes by the following equations (the bar means the arithmetic average):

$$NSE = 1 - \frac{\sum_{t=1}^{M}(O-E)^2}{\sum_{t=1}^{M}(O-\bar{O})^2}, \qquad \text{with } t = 1, \dots, M \tag{S.1}$$

$$R^2 = \left(\frac{\sum_{t=1}^{M}(O-\bar{O})\,(E-\bar{E})}{\sqrt{\sum_{t=1}^{M}(O-\bar{O})^2}\,\sqrt{\sum_{t=1}^{M}(E-\bar{E})^2}}\right)^2 \tag{S.2}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{M}(O-E)^2}{M}} \tag{S.3}$$

Software

The models are estimated utilizing the software *R* (R Core Team, 2013). We use the package *plm* for the PDMs (Croissant and Millo, 2008), the package *lme4* for the RCMs (Bates, 2010) and the package *lmtest* for the statistical tests (Zeileis and Hothorn, 2002). The robust standard errors after Arellano are computed using the *sandwich* package (Zeileis, 2004). The assignment of the weather stations and the aggregation of counties to (sub)-nations, we carried out using the *sqldf* package (Grothendieck, 2012). The maps are generated with the geographic information system software *Arc-GIS*.

Statistical tests

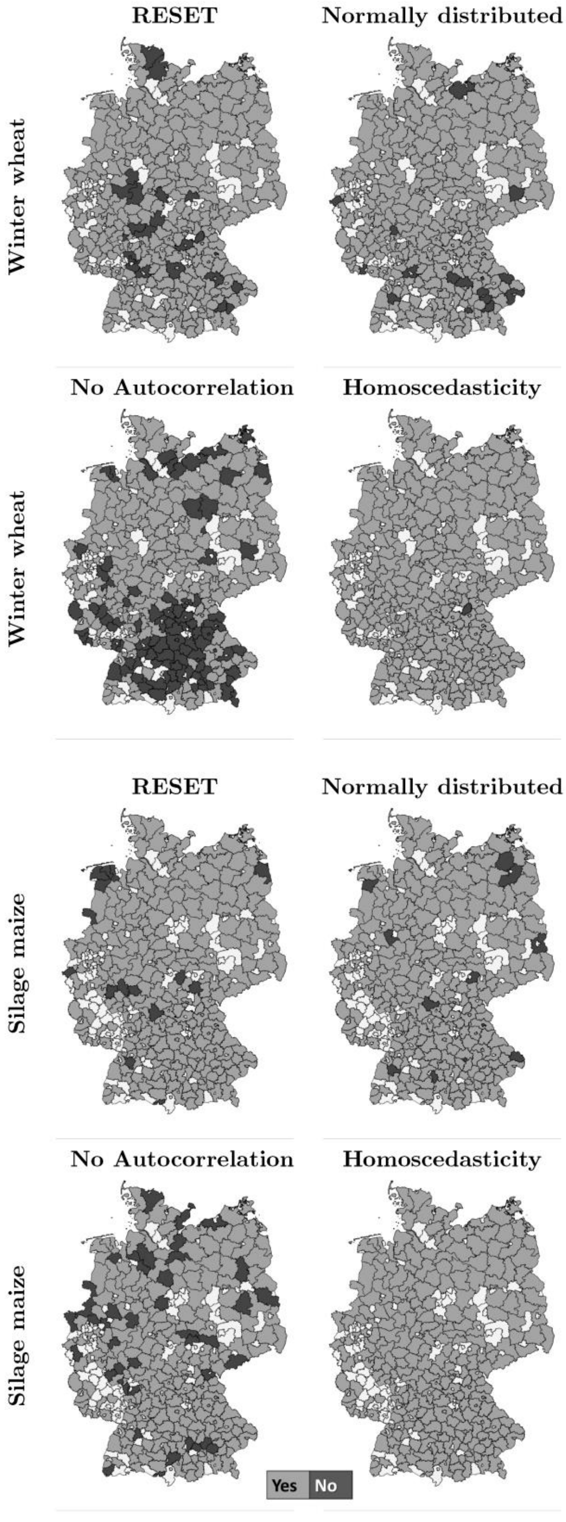


Fig. S.1: Statistical Tests of the STSMs: Regression Equation Specification Error Test (RESET, functional form): not error specified: Yes/ No; normal distributed (Shapiro-Wilk-Test): Yes/ No; no autocorrelation (Breusch-Godfrey/ Wooldridge-test): Yes/ No; homoscedasticity (Breusch-Pagan test): Yes/ No. The statistical tests are carried out with the R package *lmtest* (Zeileis and Hothorn, 2002).

Tab. S.1: Statistical Tests of the PDMs: The following statistical tests are binary coded p ≤ 0.01 → ✓, *p* > 0.01 → ×: FF (functional form): RESET Regression Equation Specification Error Test (RESET), ✓ = error specified, squared components have a significant effect; LM: Lagrange-Multiplier-test, ✓ = significant differences across counties; BG: Breusch-Godfrey/ Wooldridge-test ✓ = autocorrelation; BP: Breusch-Pagan test of heteroscedasticity, ✓ = heteroscedasticity. The statistical tests are carried out with the R package lmtest (Zeileis and Hothorn, 2002).

(Sub)Nation	FF	LM	BG	BP
Winter wheat				
Schleswig-Holstein	×	✓	✓	✓
Lower Saxony	×	✓	✓	✓
North Rhine-Westphalia	✓	✓	✓	✓
Hesse	×	✓	✓	✓
Rhineland-Palatinate	×	✓	✓	✓
Baden-Württemberg	×	✓	✓	×
Bavaria	✓	✓	✓	×
Saarland	×	×	✓	×
Brandenburg	✓	✓	✓	×
Mecklenburg-W. Pomerania	✓	✓	✓	✓
Saxony	×	✓	✓	×
Saxony-Anhalt	×	✓	✓	✓
Thuringia	✓	✓	✓	✓
Schlei/ Trave	×	✓	✓	×
Elbe	✓	✓	✓	✓
Weser	×	✓	✓	✓
Ems	×	✓	✓	×
Rhine	×	✓	✓	✓
Maas	✓	✓	×	×
Danube	×	✓	✓	×
Warnow/ Peene	✓	✓	✓	✓
Oder	×	✓	×	×
Germany	✓	✓	✓	✓
Silage maize				
Schleswig-Holstein	×	✓	✓	×
Lower Saxony	×	✓	✓	×
North Rhine-Westphalia	×	✓	✓	×
Hesse	×	✓	✓	×
Rhineland-Palatinate	✓	✓	✓	✓
Baden-Württemberg	✓	✓	✓	×
Bavaria	✓	✓	✓	✓
Saarland	✓	×	×	✓
Brandenburg	✓	✓	✓	×
Mecklenburg-W. Pomerania	✓	✓	✓	×
Sachsen	✓	×	✓	×
Saxony-Anhalt	✓	✓	✓	×
Thuringia	✓	✓	✓	×
Schlei/ Trave	×	✓	✓	×
Elbe	✓	✓	✓	✓
Weser	×	✓	✓	✓
Ems	×	✓	✓	×
Rhine	×	✓	✓	×
Maas	×	✓	×	×
Danube	✓	✓	✓	×
Warnow/ Peene	✓	✓	✓	×
Oder	×	✓	×	×
Germany	✓	✓	✓	✓

Tab. S.2: Correlation (Pearson) coefficients of the variables. The acronyms are: precipitation – PREC, potential evapotranspiration – ETP, temperature normalized solar radiation – SRT, solar radiation – R_S, fertilizer price – Fert, and acreage winter wheat – Ac WW. The month behind the variables are the corresponding period.

Variable	PREC	ETP	PREC	ETP	PREC	ETP	R_S	SRT	Fert	Ac WW
Period	May–Jul		Nov–Apr		Aug–Oct		May–Jul			
PREC May–Jul	1.00									
ETP May–Jul	-0.13	1.00								
PREC Nov–Apr	0.25	0.05	1.00							
ETP Nov–Apr	0.38	0.44	-0.04	1.00						
PREC Aug–Oct	0.34	-0.06	0.45	0.03	1.00					
ETP Aug–Oct	-0.06	0.61	-0.09	0.46	-0.39	1.00				
R_S May–Jul	-0.20	0.48	0.10	-0.13	-0.04	0.19	1.00			
SRT May–Jul	-0.16	0.14	0.09	-0.31	-0.06	0.04	0.87	1.00		
Fert	0.04	0.18	-0.06	0.17	-0.01	-0.05	0.07	0.02	1.00	
Ac WW	0.03	0.23	-0.02	0.18	0.10	0.05	0.06	-0.03	0.72	1.00

Further description of the parameters

The *SRT* May-July north-south parameter gradient for wheat reflects a similar gradient of the absolute *SRT* levels. The sensitivity is often high (low) in regions with low (high) absolute values. However, this relationship also has exceptions. The sensitivity is high and the relationship directly proportional at the coast line where the absolute level of values there is high as well. The higher responsiveness of wheat to *SRT* compared with silage maize is also reflected in the spatial patterns. The north south decline is not only weaker for silage maize than for winter wheat, but the parameter values even reverse in the east. The reason for this is unclear. A higher frequency of late frost events, which could be related to increased *SRT* values, seems to be a reasonable speculation.

Fig. S.2 depicts the yield impact of inter-annual changes in precipitation (*PREC*) from May to July for winter wheat determined from STSMs. The x-axis gives values of the precipitation parameter (β_{PREC}) normalized to the bulk of other parameters β_j (Eq. S.4). To make the positive and negative parameter values comparable, we use the absolute parameter values (hereafter expressed as sensitivity portion of *PREC* May-Jul). The y-axis shows explained yield variability of precipitation ($\beta_{PREC}\ \log x'_{PREC\ t}$) normalized to the total explained yield variability (Eq. S.5). This term is expressed as variance portion of *PREC* May-Jul. The term var is the variance.

$$\text{sensitivity portion of } PREC \text{ May} - \text{Jul} = |\beta_{PREC}| \left(\sum_{j=1}^{J} |\beta_j| \right)^{-1} \tag{S.4}$$

$$\text{variance portion of } PREC \text{ May} - \text{Jul} = \text{var}\,(\beta_{PREC}\ \log x'_{PREC\ t}) \left(var \sum_{j=1}^{J} \beta_j\ \log x'_{jt} \right)^{-1} \tag{S.5}$$

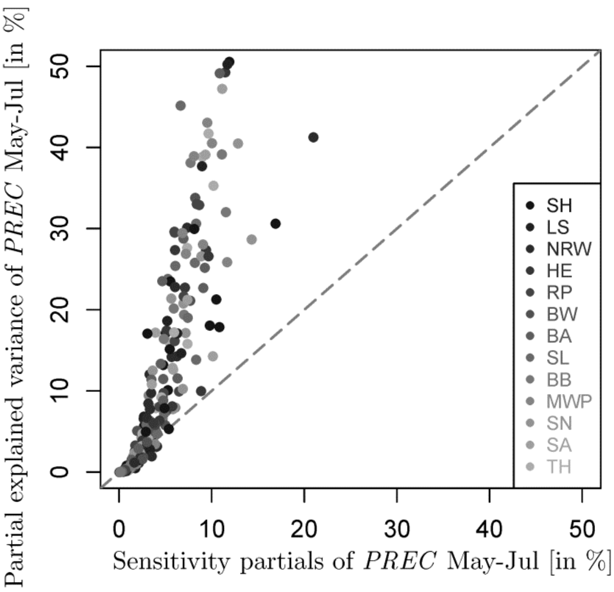


Fig. S.2: Partial explained variance of *PREC* May-Jul and sensitivity partials of *PREC* May-Jul for winter wheat determined for the STSMs. The acronyms are: SH: Schleswig-Holstein, LS: Lower Saxony, NRW: North Rhine-Westphalia, HE: Hesse, RP: Rhineland-Palatinate, BW: Baden-Württemberg, BA: Bavaria, SL: Saarland, BB: Brandenburg, MWP: Mecklenburg-Western Pomerania, SN: Saxony, SA: Saxony-Anhalt, TH: Thuringia.

References

Bates, D., 2010. lme4: Mixed-effects modeling with R. Springer.

Croissant, Y. and Millo, G., 2008. Panel data econometrics in R: The plm package. Journal of Statistical Software, 27(2): 1-43.

DWD, 2011. Daily weather data, 1951 - 2010. German Weather Service.

Federal Statistical Office, 2013. Ackerland nach Hauptfruchtgruppen und Fruchtarten.

Grothendieck, G., 2012. sqldf: Perform SQL selects on R data frames. R-Package, 0.4-6.4.

Nuzzo, R., 2014. Scientific method: Statistical errors - p values, the 'gold standard' of statistical validity, are not as reliable as many scientists assume. Nature Climate Change, 506: 150-152.

Prost, L., Makowski, D. and Jeuffroy, M.-H., 2008. Comparison of stepwise selection and Bayesian model averaging for yield gap analysis. Ecological Modelling, 219: 66-76.

R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, 3.0.0. R Foundation for Statistical Computing, Vienna, Austria.

Schrödter, H., 1985. Verdunstung: Anwendungsorientierte Meßverfahren und Bestimmungsmethoden. Springer, 204 p.

Statistical Offices of the Federation and the Länder, 2013a. Datensatz Anbaufläche (Feldfrüchte und Grünland): Deutschland, Jahre, Fruchtarten 1991-2007.

Statistical Offices of the Federation and the Länder, 2013b. Hektarerträge ausgewählter landwirtschaftlicher Feldfrüchte - Jahressumme - regionale Tiefe: Kreise und krfr. Städte.

Statistical Offices of the Federation and the Länder, 2013c. Index der Einkaufspreise landwirtschaftlicher Betriebsmittel.

Studenmund, A.H., 2000. Using Econometrics: A Practical Guide, 4. Addison Wesley.

Whittingham, M.J., Stephens, P.A., Bradbury, R.B. and Freckleton, R.P., 2006. Why do we still use stepwise modelling in ecology and behaviour? Journal of Animal Ecology, 75: 1182–1189.

Wooldridge, J.M., 2013. Introductory Econometrics. A Modern Approach. South Western Cengage Learning, 868 pp.

Zeileis, A., 2004. Econometric Computing with HC and HAC Covariance Matrix Estimators. Journal of Statistical Software, 11(10): 1-17.

Zeileis, A. and Hothorn, T., 2002. Diagnostic Checking in Regression Relationships.