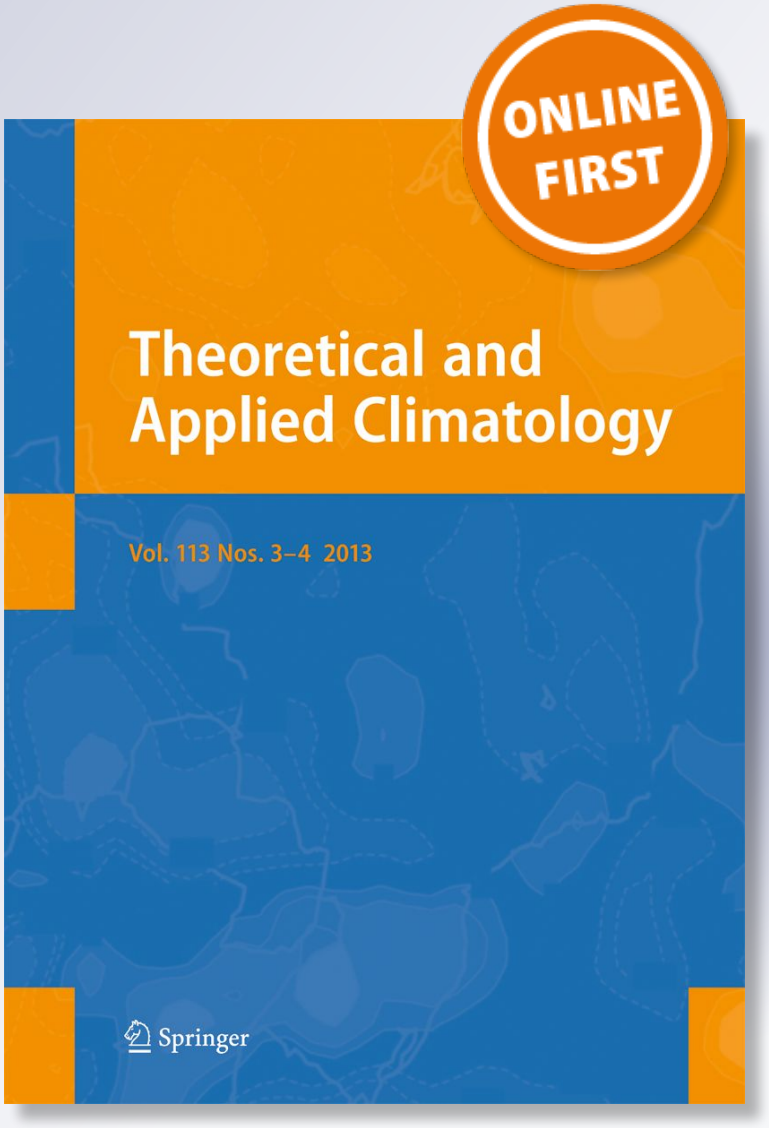
A framework for standardized calculation of weather indices in Germany

Markus Möller, Juliane Doms, Henning Gerstmann & Til Feike

Theoretical and Applied Climatology

ISSN 0177-798X

Theor Appl Climatol
DOI 10.1007/s00704-018-2473-x



Springer

Theoretical and Applied Climatology
https://doi.org/10.1007/s00704-018-2473-x

ORIGINAL PAPER

A framework for standardized calculation of weather indices in Germany

Markus Möller[1] · Juliane Doms[2] · Henning Gerstmann[3] · Til Feike[1]

Received: 23 October 2017 / Accepted: 29 March 2018

Abstract
Climate change has been recognized as a main driver in the increasing occurrence of extreme weather. Weather indices (WIs) are used to assess extreme weather conditions regarding its impact on crop yields. Designing WIs is challenging, since complex and dynamic crop-climate relationships have to be considered. As a consequence, geodata for WI calculations have to represent both the spatio-temporal dynamic of crop development and corresponding weather conditions. In this study, we introduce a WI design framework for Germany, which is based on public and open raster data of long-term spatio-temporal availability. The operational process chain enables the dynamic and automatic definition of relevant phenological phases for the main cultivated crops in Germany. Within the temporal bounds, WIs can be calculated for any year and test site in Germany in a reproducible and transparent manner. The workflow is demonstrated on the example of a simple cumulative rainfall index for the phenological phase *shooting* of winter wheat using 16 test sites and the period between 1994 and 2014. Compared to station-based approaches, the major advantage of our approach is the possibility to design spatial WIs based on raster data characterized by accuracy metrics. Raster data and WIs, which fulfill data quality standards, can contribute to an increased acceptance and farmers' trust in WI products for crop yield modeling or weather index-based insurances (WIIs).

1 Introduction

Climate change has been recognized as a main driver in the increasing occurrence of extreme weather (Field et al. 2012). Weather indices (WIs) are used to assess extreme weather conditions regarding its impact on crop yields (Goodwin and Mahul 2004; Barnett and Mahul 2007; World Bank 2011). WIs can be objectively measured during crop growth, which is not necessarily the entire growing season. Especially for yield formation, a specific phenological phase is often of major importance (Pietola et al. 2011; Castañeda-Vera et al. 2014; Vijaya Kumar et al. 2016). WIs may be based on a single weather variable or a combination of different weather variables. As such, different WIs are described in literature for applications in different regions. For instance, rainfall and drought indices have been applied for wheat in Australia (Adeyinka et al. 2016), for cereals in Morocco (Skees et al. 2001; Stoppa and Hess 2003), for wheat and rice in Nepal (Poudel et al. 2016), for maize in China (Chen et al. 2017), for cereals in West Africa (Okpara et al. 2017), or for tomato in Spain (Castañeda-Vera et al. 2014). Other successfully applied WIs are heat indices for wheat in India (Vijaya Kumar et al. 2016) and for cereals in China (Zhang et al. 2017), and frost indices for wheat in Finland (Pietola et al. 2011) and India (Vijaya Kumar et al. 2016).

Designing WIs is challenging, since complex crop-climate relationships have to be considered (Lüttger and Feike 2018). As a consequence, corresponding input data have to reflect the "temporal and spatial behavior of entities

✉ Markus Möller
markus.moeller@julius-kuehn.de

Juliane Doms
juliane.doms@landw.uni-halle.de

Henning Gerstmann
henning.gerstmann@geo.uni-halle.de

Til Feike
til.feike@julius-kuehn.de

[1] Julius Kühn Institute (JKI) - Federal Research Centre for Cultivated Plants, Institute for Strategies and Technology Assessment, Stahnsdorfer Damm 81, 14532 Kleinmachnow, Germany

[2] Martin Luther University Halle-Wittenberg, Institute of Agricultural and Nutritional Sciences, Agribusiness Management Group, Karl-Freiherr-von-Fritsch-Str. 4, 06120 Halle (Saale), Germany

[3] Martin Luther University Halle-Wittenberg, Institute of Geosciences and Geography, Von-Seckendorff-Platz 4, 06120 Halle (Saale), Germany

Published online: 04 May 2018

and local conditions" (Lokers et al. 2016). Especially in the context of designing weather index-based insurances (WIIs), data quality issues have been identified, which can contribute to imperfect correlations of the experienced loss and the corresponding WI. Apart from the fact that a specific WI might be an inappropriate indicator for yield losses, such poor correlations are often related to the spatio-temporal quality of input data for designing WIs:

- Temporal inaccuracies can arise when a WI does not match periods of crops' sensitivity to specific (harmful) weather conditions (Dalhaus et al. 2018). This problem is often relevant when fixed start and end calendar dates are used for the definition of reference periods (e.g., Turvey 2001, Pelka and Musshoff 2013). Therefore, a promising strategy to tackle the temporal inaccuracy problem is the consideration of phenological information. In doing so, relevant time frames can be determined in a flexible manner reflecting the inter-annual variability of relevant phenological phases of plant development. For instance, Conradt et al. (2015) could show in a WII design study in a Kazahk test site how start and end dates of time frames can be defined by using *growing degree days* (GDD). This approach assumes strong relations between specific temperature sums and phenological stages (Chuine et al. 2003). As shown in Dalhaus et al. 2018, an alternative to GDD is the consideration of phenological observations, which are monitored in some countries by public institutions (Schwartz 2006).
- Spatial inaccuracies can be caused by differences between site-specific weather or phenological conditions and the point of measurement or observation (Grassini et al. 2015). Thus, Dalhaus and Finger (2016) suggest the usage of raster data for WI calculation. Apart from advantages of such data for practical implementations, WIs could be calculated without the consideration of spatial and temporal data gaps.

It becomes obvious that data of high spatio-temporal quality data are crucial for designing WIs. In the WII context, Leblois and Quiron (2013) emphasize that insurance adoption and market acceptance are related to the farmers' trust in the WI products as well as in the supplying institutions. Consequently, Dalhaus and Finger (2016) conclude that "a straightforward WI calculation [should be] based on data provided by a trustworthy public institution".

Institutionally, provided geodata have to meet quality requirements of ISO standards and standards for geospatial web services. ISO standards represent norms for cartographic products that help geodata producers objectively determine and describe the quality of geodata using statistic metrics (Möller et al. 2013; Lokers et al. 2016). Quality metrics for geospatial web services also address the dynamic aspect of geodata accuracy "enabling [them] to be used in geoprocessing procedures or in decision-making with maximum reliability" (Cruz et al. 2012).

In this article, we present a WI design framework for the total area of Germany, which enables the automatic and operational derivation of standardized and dynamic WIs, where all steps of geodata processing are reproducible and the geodata products are characterized by spatial accuracy metrics. The framework is explained on the example of a simple cumulative rainfall WI, which is based on the combination of phenological and precipitation raster data sets (Rauthe et al. 2013; Gerstmann et al. 2016). The WI is calculated for 16 German test sites, the period between 1994 and 2014 and the phenological phase *shooting* of winter wheat due to its importance for harvest yield (Acevedo et al. 2002; Conradt et al. 2015). To highlight the potential of the novel approach for WI standardization, we finally compare the resulting spatial WI with a station-based WI variant.

2 Materials and methodology

2.1 Test sites

For the analysis, 16 test sites are investigated located in the German federal states North Rhine-Westphalia and Lower Saxony (Fig. 1a). The climatic situation is classified as warm, humid, and continental. In North Rhine-Westphalia, the average annual temperature is 8.9 °C, and in Lower Saxony 8.6 °C. The average annual precipitation is 875 mm in North Rhine-Westphalia and 746 mm in Lower Saxony. The test sites were selected as they are part of an analysis, which investigates the suitability of precipitation-based WIIs for risk reduction in regions with moderate climatic conditions (Doms et al. 2017, 2018).

2.2 Workflow

In Fig. 2, the principle workflow for the standardized and automatic WI calculation based on raster data (WI^{R}) is shown, which is implemented within the statistical computing environment `R` (R Core Team 2017). The WI design concept applied in this study is based on an index variant introduced by (Dalhaus and Finger 2016), which was designed as drought risk indicator for winter wheat based on public and open data. Using the Germany-wide raster data set REGNIE of daily precipitation (Sections 2.2.1 and 2.3.2), they calculated the precipitation sum $\sum(P^{RU,Y})$ for a specific year (Y) and reference unit (RU; for example farms or test sites) during the phenological phase *shooting* of winter wheat (Eq. 1). The phase's start and end dates $DOY^{A,B}$ were determined by using phenological observations (Section 2.3.2), which were aggregated to phenological regions (Ssymank 1994).

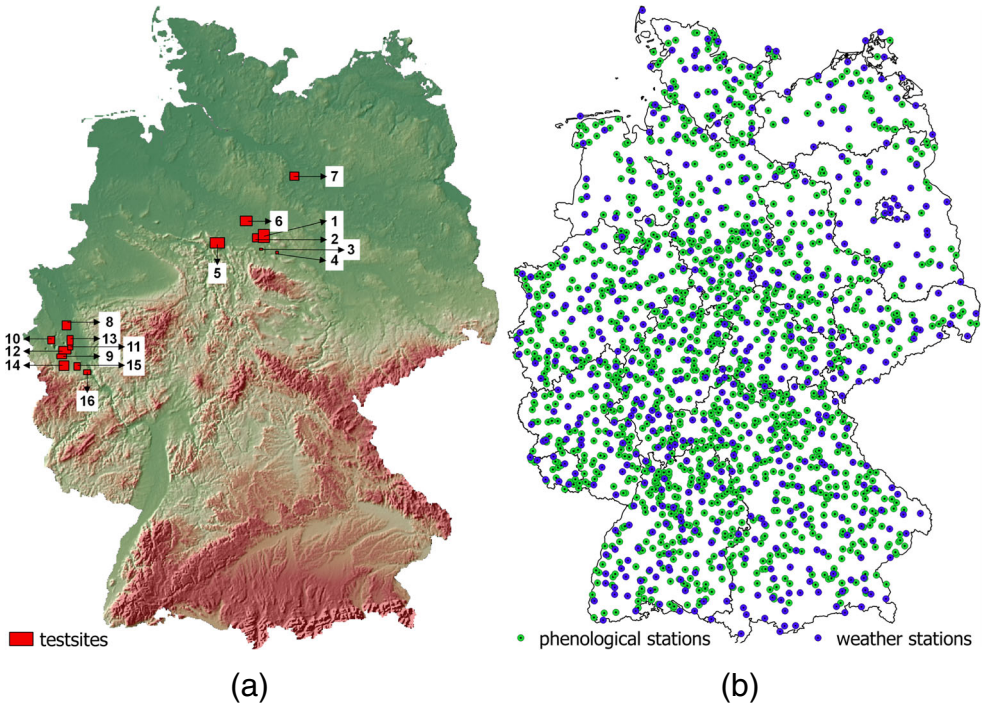


Fig. 1 The locations of test sites (**a**) as well as of phenological and meteorological stations (**b**)

In this study, the phenological model PHASE has been applied to generate phenological raster data sets. They can be used for an automatic and dynamic determination of phenological windows ($DOY^{A,B}$; Section 2.2.2) and WI derivation.

$$WI^{\mathrm{R}} = \sum_{DOY^{A}}^{DOY^{B}} (P^{RU,Y}) \tag{1}$$

A station-based WI variant (WI^{S}) has been derived using the nearest stations of meteorological measurements and phenological observations, for which temporally complete data sets were available (Section 2.3.3). The differences and relations between precipitation sums based on station data and raster data has been investigated by applying a correlation analysis as well as linear and non-linear regression (Section 2.2.3).

2.2.1 Daily precipitation and weather index

Germany-wide 1 × 1 km raster data sets of daily precipitation (Section 2.3.2) have been derived by using the `R` function `read_regnie`, which is part of the package `esmisc` (Szoecs 2016). The actual test site-specific calculation of

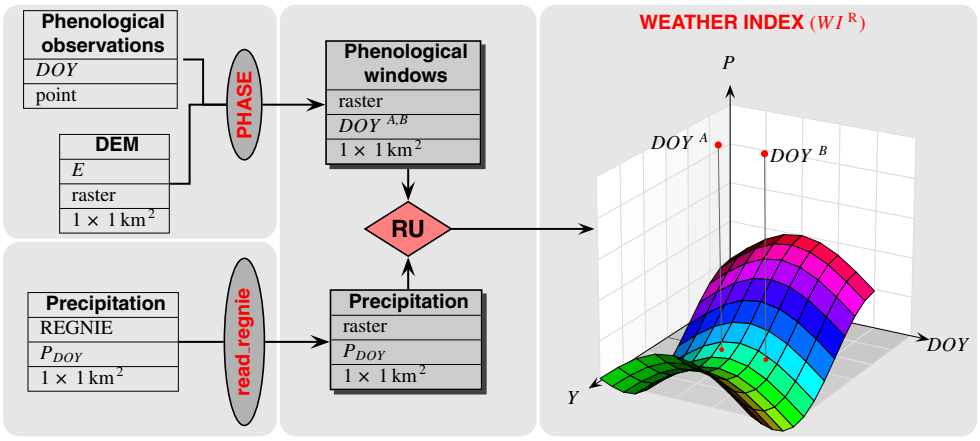


Fig. 2 Principle workflow for the derivation of a phase-specific and raster-based weather index WI^{R}. RU – reference unit | DOY – phenological day of the year | P_{DOY} – daily precipitation | $DOY^{A,B}$ – DOYs of a phase's start and end | DEM – digital elevation model | E – elevation | Y – year | PHASE – model for the interpolation of beginning phenological phases | read_regnie – model for the import of REGNIE data

the raster-based WI has been realized by zonal statistics functions of the R package `raster` (Hijmans 2016).

2.2.2 Phenological phases and windows

The PHASE model has been applied to interpolate Germany-wide raster data of phase-specific DOYs (Gerstmann et al. 2016). The model is based on the growing degree days concept, which relates phenological events of plants' development to phase-specific accumulated heat sums. The model approach is based on the assumption that temperature is the main driving factor for intra-seasonal timing of phenological events in temperate regions like Central Europe (Chmielewski et al. 2004). During the modeling procedure, an *indicator temperature sum* is determined by analyzing the distribution of the temperature sums accumulated between sowing and the date of phase observation. Germany-wide temperatures result from the interpolation of daily mean temperatures, which are provided by DWD weather stations (Section 2.3.2). For each location, the day, on which the accumulated temperature sum exceeds the *indicator temperature sum*, is modeled and spatially interpolated using regression kriging (Hengl et al. 2007). In addition, the accuracy metric *kriging standard deviation* (σ^K) is calculated for each phenological interpolation result, which represents the spatial standard error of estimation computed for a kriged estimate (Hiemstra et al. 2009).

Following Möeller et al. (2017), the interpolation results can be used for the definition of the phenological phase *shooting* which represents the temporal window between the beginning of the phases *shooting* (WW^{15}) and *heading* (WW^{18}). All interpolations consider digital elevation data (Section 2.3.1).

2.2.3 Correlation analysis, linear and non-linear regression

Differences between WIs derived form station-based and raster data have been analyzed by applying correlation and regression functions:

- Using the basic function `cor.test`, the non-parametric *Spearman's rank correlation* results in the coefficient ρ, which is an indicator for the similarity between two sets of ranked variables (Davis 2002).
- Both a linear regression model and the non-linear *Random Forest* model are included in the R package `caret` (Kuhn et al. 2014). The function `train` formalizes the training, pre-processing, tuning, and performance assessment of a wide variety of spatial modeling techniques. In order to avoid over-fitting, each model is based on bootstrapped training samples (25 iterations). The modeling performances are assessed by the metrics *root mean square error* ($RMSE$) and *coefficient of determination* (R^2) based on observed and predicted values (Kuhn 2008; Kuhn and Johnson 2013).

Apart from statistical accuracy metrics, the modeling results are evaluated according to their relationship to each predictor. The resulting variable importance for the linear regression model corresponds to the absolute value of the t-statistic for each model parameter. For *Random Forest* regression, the importance of each explaining variable is derived from the percent increase in *mean squared error* (MSE), which results from the permutation of *out-of-bag* data for each variable. Finally, the variable importance metrics are scaled between 0 and 100.

2.3 Data

2.3.1 SRTM DEM

The Shuttle Radar Topography Mission (SRTM[1]) resulted in a public and open digital elevation model (DEM) for almost the entire Earth except the polar regions. It has a geometric resolution of $\approx$ 30 $\times$ 30 m, with horizontal and vertical accuracies of about 20 and 16 m, respectively (Rabus et al. 2003). A Germany-wide DEM has been filtered (Lee 1980) to reduce signal noise and then aggregated to 1 $\times$ 1 km raster size with a total number of 358,320 pixels. The raster size roughly corresponds to the positional inaccuracy of the phenological observations (Section 2.3.2). Figure 1a shows the resulting DEM as colored hill shade for the total area of Germany.

2.3.2 Phenological and meteorological data

In Germany, a phenological and meteorological monitoring network is driven by the German Weather Service (DWD; in German: *Deutscher Wetterdienst*). The phenological network of yearly observers consists of approximately 1200 volunteers, which map the beginning of principle phenological phases of the most frequently cultivated crops according to standardized criteria since 1951 (Kaspar et al. 2014). Each plant is observed on a different number of stations, depending on the abundance and agrometeorological relevance of the respective crop type. The mapping results are reported at the end of the year under review and are checked for plausibility. The positional accuracy of the point data set is about 2 $\times$ 2 km. The reported crop types and corresponding phenological phases are listed in Kaspar et al. (2014). In this study, the beginning phases *shooting* and *heading* of winter wheat are considered. Daily meteorological measurements (mean

[1] Link to USGS web portal EarthExplorer for downloading SRTM data: https://earthexplorer.usgs.gov

temperature) of approximately 500 stations were used, which are also provided by the DWD. Both phenological and meteorological data can be accessed via FTP.[2]

Furthermore, the DWD enables open access to daily precipitation data, which are calculated using the specific regionalization method REGNIE (Rauthe et al. 2013). Daily precipitation data between 1931 and today measured at ≈ 2000 weather stations irregularly distributed over Germany are interpolated on a raster grid of 1 × 1 km. The REGNIE procedure combines a multiple linear regression including different influencing factors such as latitude and longitude of a station and inverse distance weights. REGNIE data can also be directly downloaded from the FTP server of the DWD[3]. They are stored in a specific ASCII format, which requires a transformation into a common raster format (Section 2.2.1).

2.3.3 Distances between test sites and weather stations or phenological observations

Table 1 lists the linear distances between the centroids of the test sites as well as the nearest meteorological and phenological stations, for which a complete temporal data set between 1994 and 2014 is available. In total, data of twelve weather stations are used, five located in North Rhine-Westphalia and seven in Lower Saxony. Due to the temporal discontinuity of the phenological observations, only five stations are available providing data for the complete observation period. Three stations are located in North Rhine-Westphalia and two in Lower Saxony. Since the stations are distributed irregularly (Fig. 1b), the distances between test sites and weather stations or phenological observations fluctuate between 2 and 26 km and 0 and 90 km, respectively.

Table 1 Distances [km] between test sites and weather stations (D^{WS}) or phenological observations (D^{PS}) as well Spearman correlation coefficients (ρ) between test site-specific weather indices based on station (WI^S) or raster data (WI^S)

TS	D^{WS}	D^{PS}	$\rho(WI^S, WI^R)$
1	11	11	0.85
2	5	11	0.86
3	10	18	0.87
4	12	27	0.90
5	8	32	0.22
6	2	15	0.39
7	13	90	0.90
8	13	30	0.45
9	5	18	0.51
10	21	29	0.28
11	21	0	0.37
12	21	14	0.34
13	26	5	0.36
14	23	0	0.71
15	8	11	0.83
16	8	20	0.87

3 Results

3.1 Phenological patterns

Figure 3 shows the interpolation results of the beginning phenological phases *shooting* and *heading* of winter wheat in 2007 and 2013. These years are chosen because they are an example for the inter-annual DOY variability of phenological phases. Although the Germany-wide phenological patterns are similar, the figure reveals that the phases *shooting* and *heading* started much later in 2013 compared to 2007. With regard to spatio-temporal differences, DOY patterns show similarities with earlier dates of phase occurrence in favored regions especially in the southwest, west as well as northeast of Germany and delayed plant development in more mountainous regions.

Figure 4 summarizes all DOYs of both beginning phenological phases between 1994 and 2014. The test site-specific perspective in Fig. 4a illustrates that both phases start usually earlier in North Rhine-Westphalia (test sites 8 to 16). The later beginning of the two phases in Lower Saxony (test sites 1 to 7) is related to the increasing continental climatic influence and consecutive lower average temperatures. The average DOY values show a rather low variation between the test sites, while also the magnitude of inter-annual variation is rather constant over the sites. The year-specific perspective (Fig. 4b) reveals sometimes large inter-annual differences in average DOYs, e.g., 2007 vs. 2013, while also the inter-site variability can vary strongly between different years, e.g., 1995 vs. 1996.

The raster data sets of the beginning phases can be used to derive phenological windows for specific years and test sites based on all interpolated phenological events. This is shown in Fig. 5 on the example of test site 1 for 2007 and 2013. In doing so, test site-specific differences in phenological timing can be illustrated.

[2]Link to the DWD climate data server for downloading phenological and meteorological observations: ftp://ftp-cdc.dwd.de/pub/CDC/observations_germany

[3]Link to the DWD climate data server for downloading REGNIE data: ftp://ftp-cdc.dwd.de/pub/CDC/grids_germany/daily/regnie

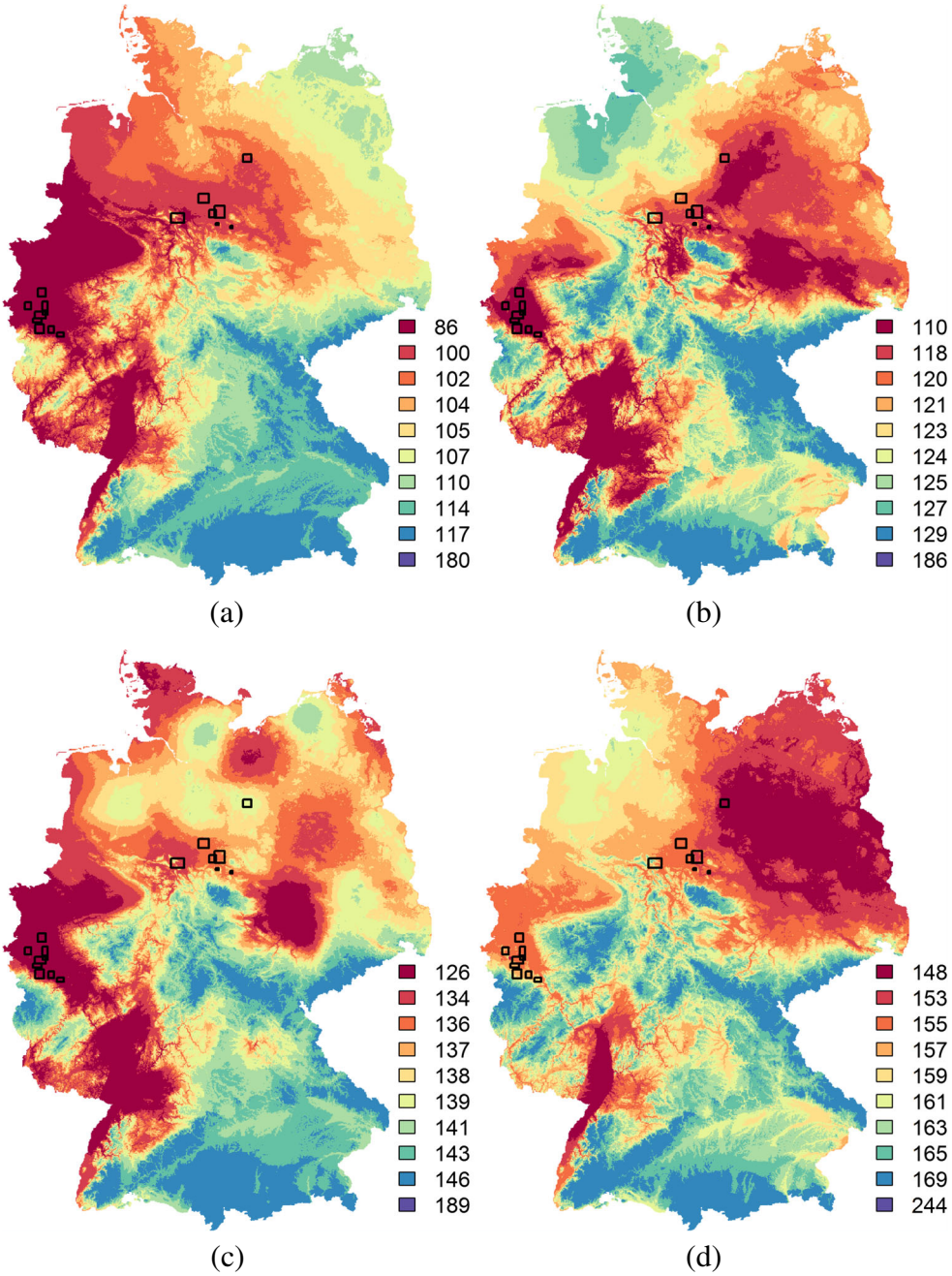


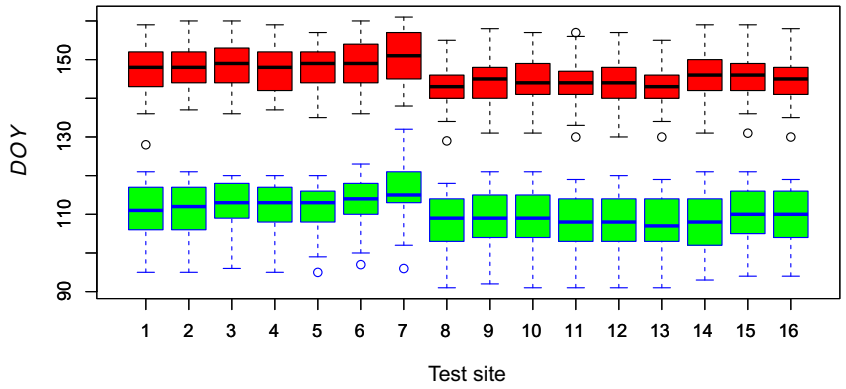
Fig. 3 Predicted phenological events (DOY) of the beginning phenological phases *shooting* (**a**, **b**) and *heading* (**c**, **d**) of winter wheat in 2007 (**a**, **c**) and 2013 (**b**, **d**). The test site names are shown Fig. 1b

3.2 Model inaccuracies

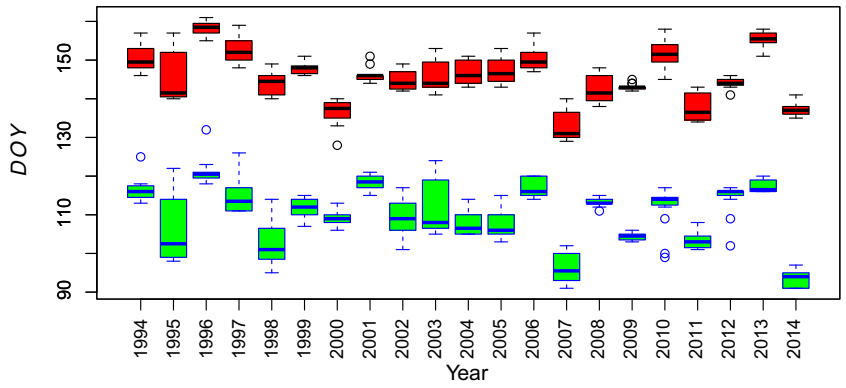
Each interpolated value of the beginning of a phenological phase is related to a corresponding spatial data set of the *kriging standard deviation* (σ^K), which is shown in Fig. 6 on the example of *shooting* and *heading* of winter wheat in 2007 and 2013. Accordingly, σ^K values vary over phases, sites, and years. Higher σ^K values prevail especially in some mountainous regions, e.g., the Alps in the south, as well as in the north and northeast where both the meteorological and phenological observation density is less than in the remaining area of Germany (see Fig. 1b).

The year- and test site-specific means of *kriging standard deviation* ($\overline{x}(\sigma^K)$) for the beginning phases *shooting* and *heading* are displayed in Fig. 7. The $\overline{x}(\sigma^K)$ values for the beginning phase *shooting* vary from 4.8 to 6.2 days and for *heading* from 2.9 to 3.9 days (25th and 75th percentiles).

The $\overline{x}(\sigma^K)$ variations of all test sites are comparable (Fig. 7a). Similar to the strong year-to-year variation in the absolute DOY values for the beginning phases (cf., Fig. 4b), there is also a strong variation in the phase-specific model inaccuracy over the years (Fig. 7b). Additionally, there are other factors mainly linked to abiotic stress like water

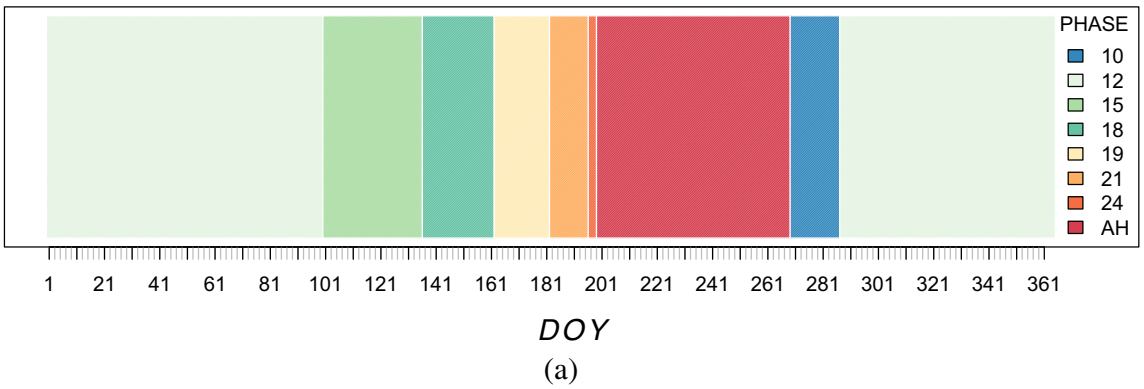


(a) Test site-specific DOY-year boxplots

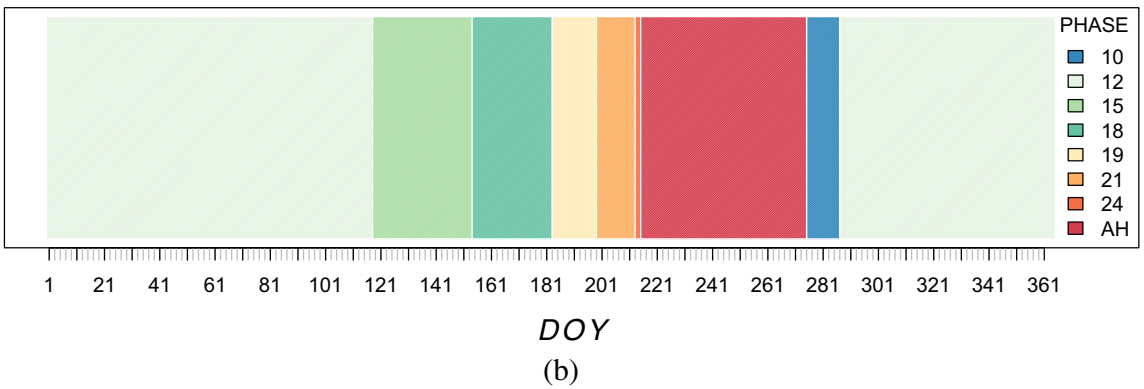


(b) Year-specific DOY-test site boxplots

Fig. 4 Test site- (**a**) year-specific DOY distributions (**b**) for the beginning phenological phases 15 (green/blue) and 18 (red/black) during the period between 1994 and 2014



(a)



(b)

Fig. 5 Phenological phases of test site 1 for 2007 (**a**) and 2013 (**b**). Phase IDs: 12 - emerging · 15 - beginning of shooting · 18 - beginning of heading · 19 - milk ripeness · 21 - yellow ripeness · 24 - harvest · AH - after harvest

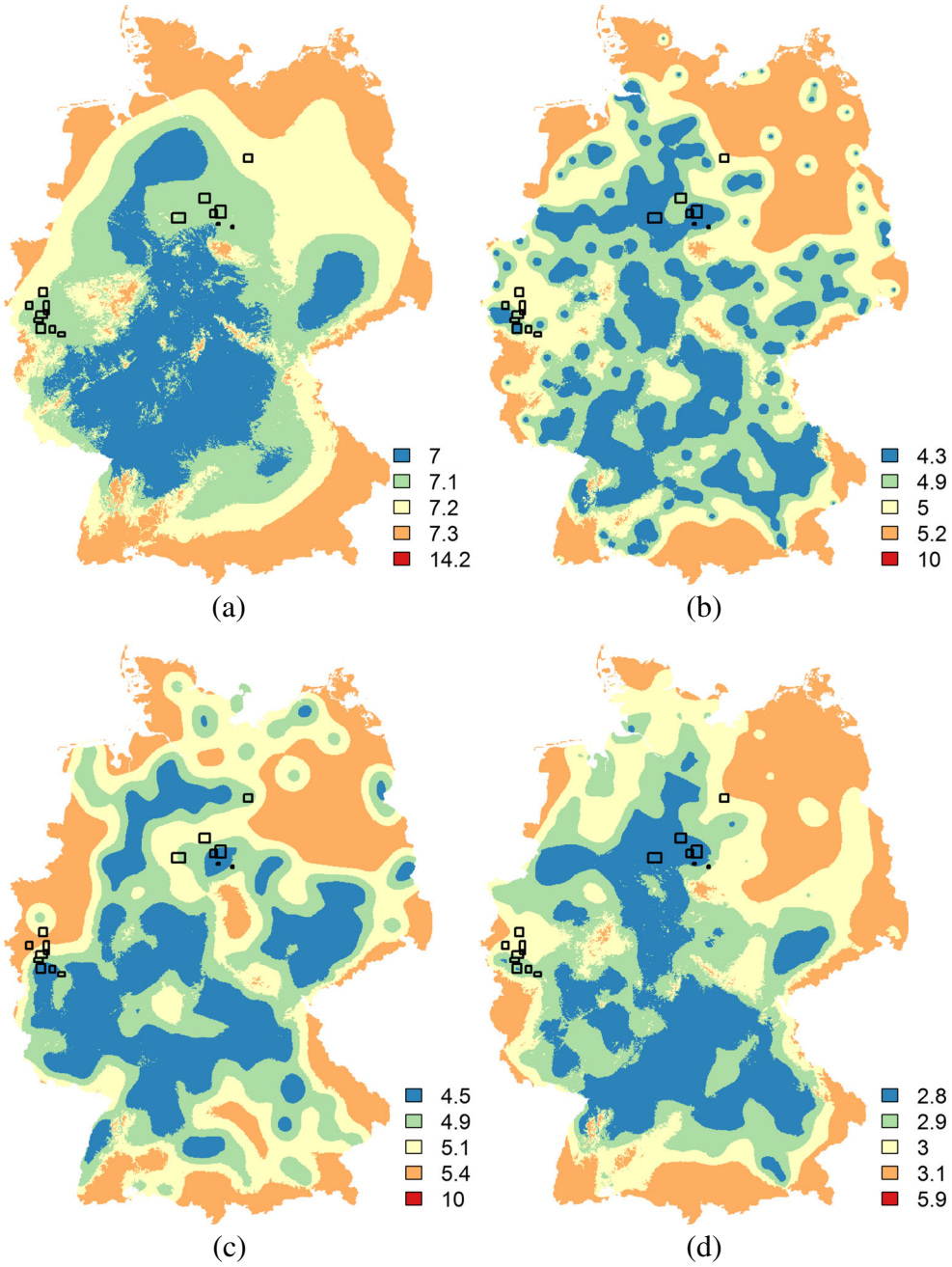


Fig. 6 *Kriging standard deviation* (σ^K) for the interpolated beginning phenological phases *shooting* (**a**, **b**) and *heading* (**c**, **d**) of winter wheat in 2007 (**a**, **c**) and 2013 (**b**, **d**). The test site names are shown Fig. 1b

availability and soil moisture, soil properties, and fertilization but also crop cultivar that affect the phenological development of winter wheat (McMaster and Wilhelm 2003; Nellis et al. 2009; Zhao et al. 2013; Rezaei et al. 2018). As such, the model inaccuracies express differences in the year-specific importance of the explaining variable *temperature sum* (Section 2.2.2).

Both boxplot diagrams reveal that the beginning phase *shooting* is characterized by higher $\overline{x}(\sigma^K)$ values compared to the later phase *heading*. Following Gerstmann et al. (2016), this behavior is correlated to the strong anthropogenic influence, namely the timing of sowing, on early phenological development. Over the growing season, the effects of differing sowing dates on phenology continuously diminish due to the increasing cumulative temperature effect, which is well captured by the PHASE model. Consequently, the model inaccuracy, expressed by the phase-specific $\overline{x}(\sigma^K)$ values, decreases.

3.3 Precipitation sums

The two Germany-wide REGNIE data sets for the DOYs 127 in 2007 and 145 in 2013 illustrate the temporal and spatial dynamic of daily precipitation (Fig. 8a, b). While

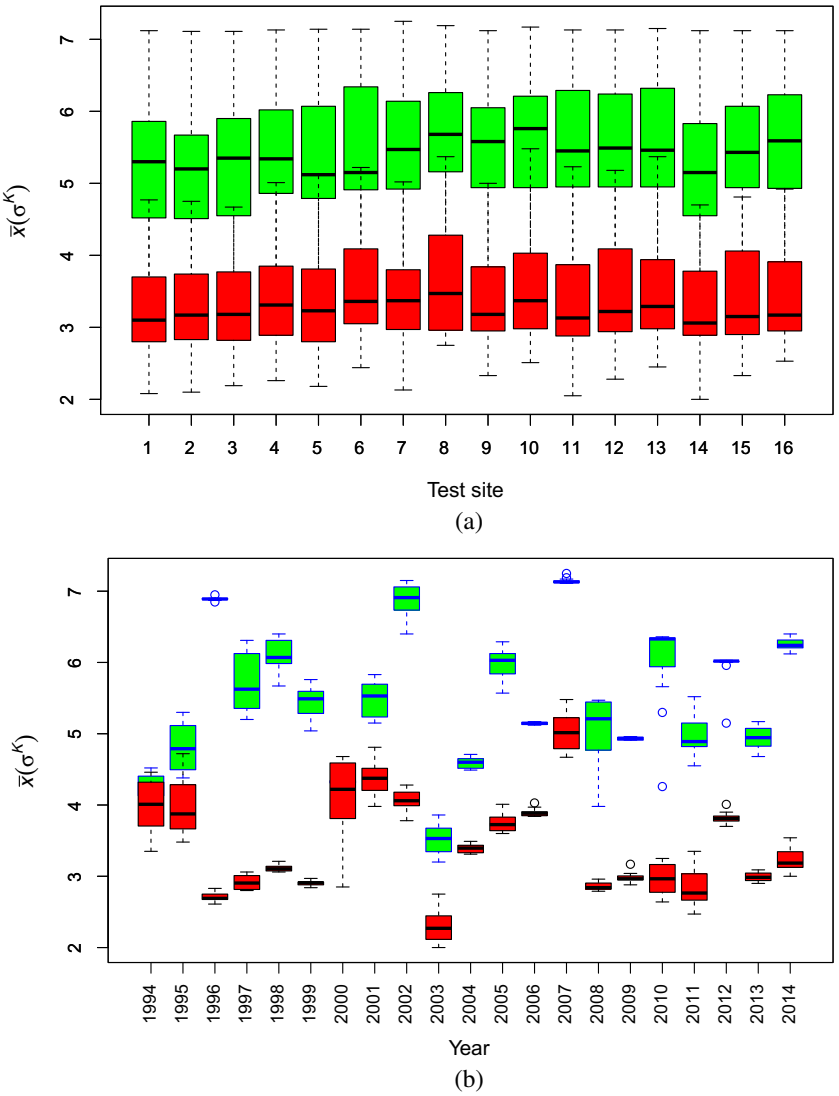


Fig. 7 Test site- (**a**) and year-specific means (**b**) of *kriging standard deviations* ($\overline{x}(\sigma^K)$) for the phases *shooting* (WW^{15}; green/blue) and *heading* (WW^{18}; red/black) during the period between 1994 and 2014

there was little rain in the North Rhine-Westphalian test sites, almost all test sites in Lower Saxony were exposed to significant precipitation. This is also true for test site 1, for which both rain events occurred during the phenological phase *shooting*. The graphs in Fig. 8c, d show phase-specific daily precipitations as well as the corresponding phase-specific precipitation sums (WI^{R}). The DOYs 127 and 145 are highlighted. While the phase duration of 36 days is identical for both years, the phases' starting and end points differ considerably. In 2007, the period begins on the $DOY = 100 \pm 7$ and ends on the $DOY = 136 \pm 5$. In 2013, the starting and end points are $DOY = 118 \pm 5$ and $DOY = 154 \pm 3$ (see Section 3.1). The given uncertainties result from the test-site-specific $\overline{x}(\sigma^K)$ values. Accordingly, 2007 was characterized by higher uncertainties than 2013 (see Fig. 7b).

Figure 8c, d also reveals that the precipitation sums are in 2013 ($WI^{R} = 188$ mm) more than twice as high as in 2007 ($WI^{R} = 73$ mm). Considering the uncertainties of phases' starting and end points, the summation results can vary between 55 and 76 mm in 2007 as well as 187 and 203 mm in 2013.

Figure 9 presents the year- and test site-specific WI^{R} distributions for the phase *shooting*. Contrary to the entrance of phenological phases shown in Fig. 4, no regional differences are visible. However, all test sites show a high variation of precipitation sums over the years (Fig. 9a). The year-specific WI^{R} distribution reveals strong inter-annual changes of precipitation sums (Fig. 9b). Both figures demonstrate that even regions with moderate climate conditions are characterized by a high spatio-temporal precipitation variability.

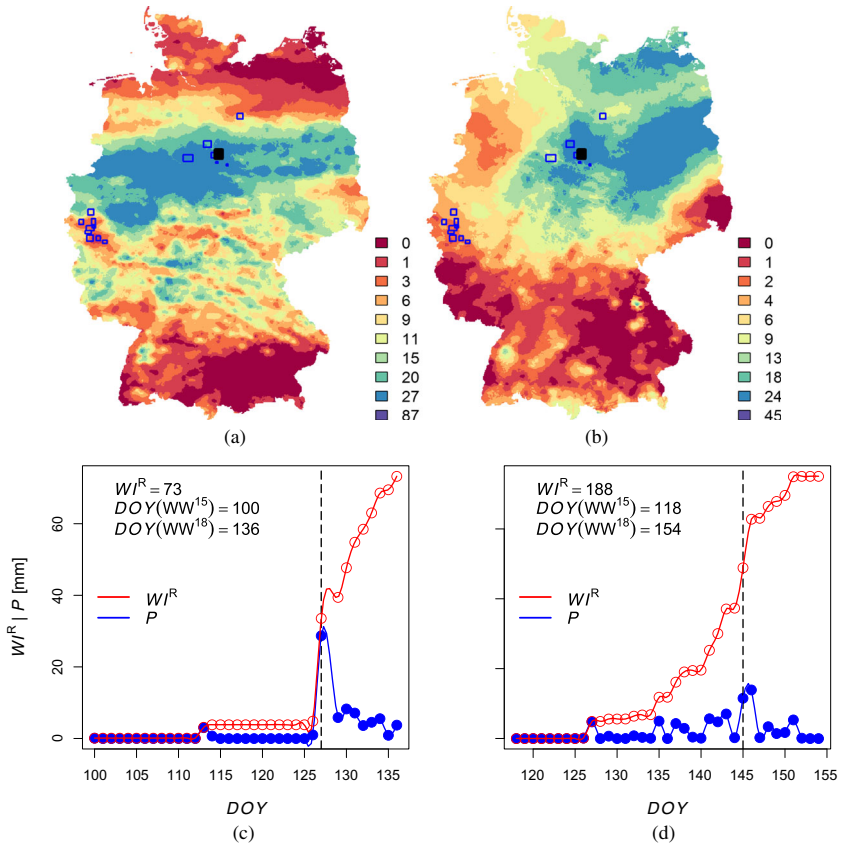


Fig. 8 Two REGNIE precipitation raster (P) [mm] for (**a**) $DOY = 127$ (May 7th 2007) and (**b**) $DOY = 145$ (May 25th 2013) as well as phase-specific precipitation and corresponding WIs (P and WI^{R} in mm) for test site 1 in 2007 (**c**) and 2013 (**d**). Test site 1 is marked black (**a**, **b**). Both DOYs are emphasized by dashed black vertical lines (**c**, **d**)

3.4 Comparison of station-based and raster-based WIs

Figure 10 summarizes test site-specific scatterplots of WIs based on stations (WI^{S}) and raster data (WI^{R}) for the phase *shooting* from 1994 to 2014. The plots indicate that some test sites (e.g., 1 to 4 or 7) are characterized by strong relations between WI^{R} and WI^{S}. Other test sites (e.g., 10 to 13) are weakly correlated. In Table 1, all corresponding *Spearman correlation* coefficients ($\rho(WI^{\mathrm{S}}, WI^{\mathrm{R}})$) are listed.

In the following, we tested if the differences between both WI variants can be explained by the measured distances between test site centroids and the locations of meteorological stations (D^{WS}) or phenological observations (D^{PS}; Table 1; see Section 2.3.3). In doing so, we set up a linear (LR) and non-linear regression model (RF) according to Eq. (2).

$$WI^{\mathrm{R}} \sim WI^{\mathrm{S}} + \overline{x}(\sigma^K)(\mathrm{WW}^{18}) + \overline{x}(\sigma^K)(\mathrm{WW}^{15}) + D^{\mathrm{WS}} + D^{\mathrm{PS}} \quad (2)$$

Apart from both distances, we also considered test site-specific means of the *kriging standard deviations* of the beginning phases *shooting* ($\overline{x}(\sigma^K)$ (WW^{18})) and *heading* ($\overline{x}(\sigma^K)$ (WW^{18})) as additional explaining variables.

Table 2 summarizes the accuracy metrics R^2 and $RMSE$ as well as the variable importances of both models. Accordingly, the variable W^S is most important for both models. However, the ranking of the remaining variables differs. In the LR model, the differences between WI^{S} and WI^{R} are also explained by the variable D^{WS}. This is in contrast to the RF model, where both $\overline{x}(\sigma^K)$ variables show higher explanatory power compared to the variable D^{WS}. Comparing the accuracy metrics $RMSE$ and R^2, the RF model clearly outperforms the LR model, which is shown in Fig. 11. There, three scatter plots are overlaid. The black colored scatterplot relates all original WI^{S} and WI^{R} values. The red and blue scatterplots illustrate the predictive power of the LR and RF model with consideration of the additional explaining variables.

4 Discussion

4.1 Towards standardized geodata for designing weather indices

The approach presented in this study can be considered as a dynamic WI design framework, which we illustrated on a simple precipitation index. The operational process chain enables the dynamic and automatic definition of temporal

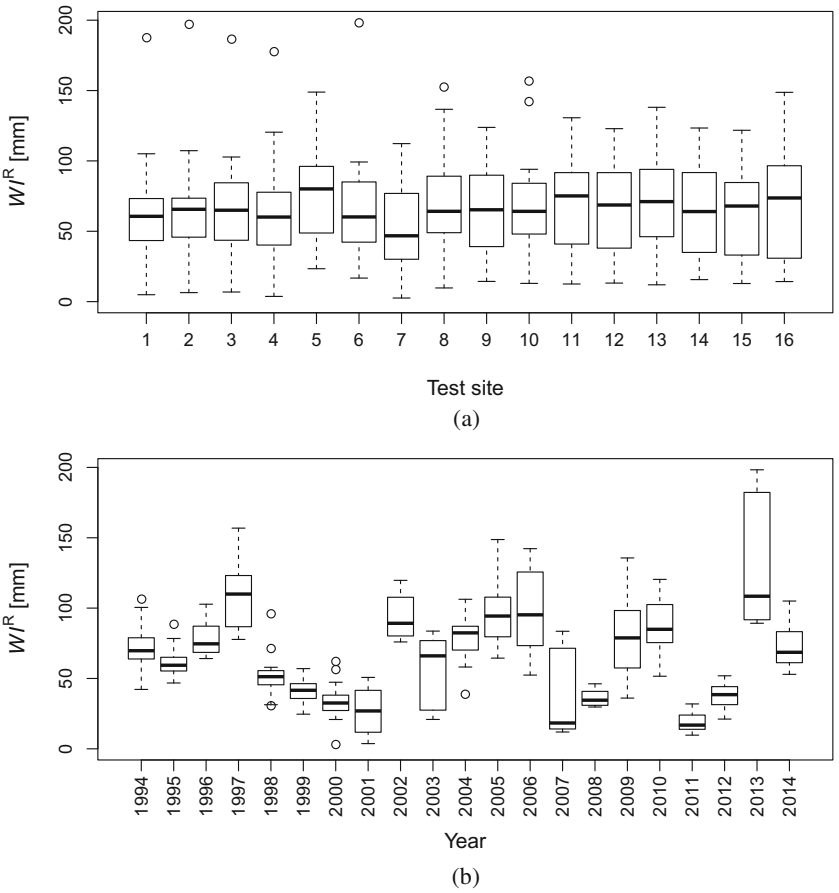


Fig. 9 Test site- (**a**) and year-specific distributions (**b**) of phase-specific WI (WI^R) [in mm] illustrated on the example of the phase *shooting* during the period between 1994 and 2014

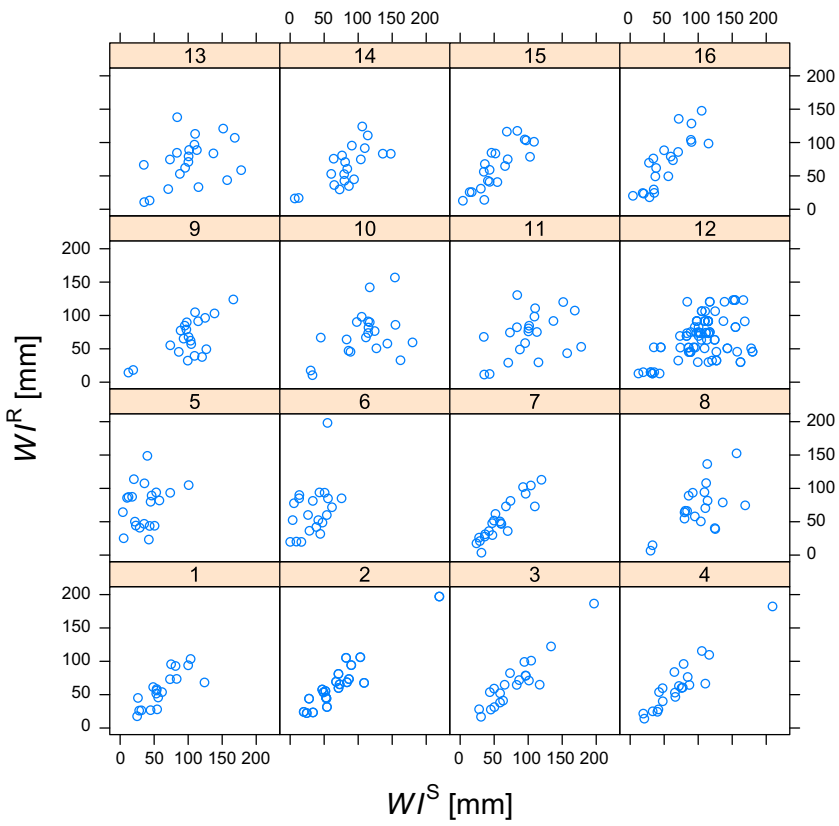


Fig. 10 Test site-specific scatter plots of phase-specific precipitation sums based on meteorological stations (WI^S) and raster data (WI^R) for the phase *shooting* from 1994 to 2014

Table 2 Variable importance and accuracy metrics for models based on linear regression (LR) and random forest (RF; see Eq. (2) and Fig. 11)

Variable	Importance	
	LR	RF
WI^{S}	100	100
$\overline{x}(\sigma^{K})$ (WW^{18})	6.64	44.18
$\overline{x}(\sigma^{K})$ (WW^{15})	7.24	39.51
D^{WS}	37.28	19.06
D^{PS}	0	0
Metric	Accuracy	
	LR	RF
$RMSE$	27.59	20.65
R^2	0.33	0.63

windows of relevant development stages of main cultivated crops in Germany. In doing so, interactions between extreme weather conditions and specifically sensitive periods of phenological development can be considered in WI design (see Conradt et al. 2015, Dalhaus and Finger 2016, Dalhaus et al. 2018). Within the temporal bounds, WI variants (e.g., Luttger and Feike 2018, Goemann et al. 2015) can be calculated for any year and test site in Germany by using open public data provided by state authorities. This means that seeking for suitable weather and phenological stations next to fields or farms becomes superfluous for

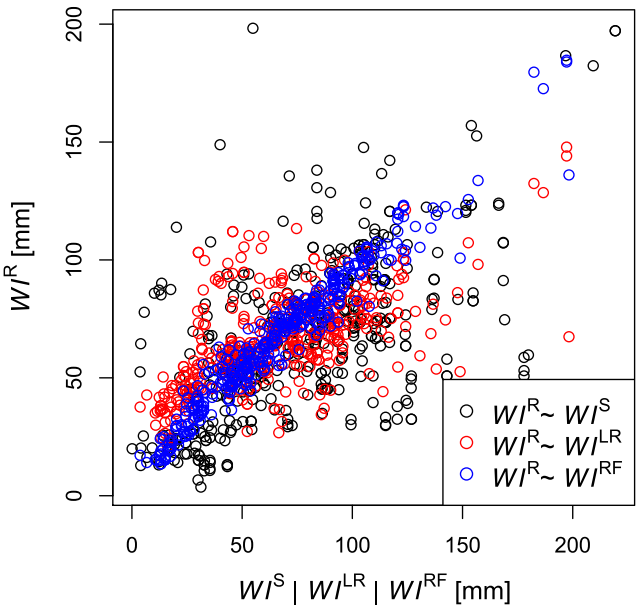


Fig. 11 Scatter plots of phase-specific precipitation sums based on raster data (WI^{R}) and weather stations or phenological observations (WI^{S}) as well as relations between WI^{R} and weather indices predicted via linear regression (WI^{LR}) or non-linear *Random Forest* regression (WI^{RF}) according to Eq. (2) (see Table 2).

WI calculation. In addition, the procedure fills spatial and temporal gaps of phenological and meteorological station data sets.

4.2 WI standardization

As shown in Section 3.4, we detected significant discrepancies between station- and raster-based WIs. The bias is related to distances between the locations of test sites and weather stations (D^{WS}) or phenological observation points (D^{PS}). Since the distances vary unsystematically over space and time, the quality of WI design approaches, building on station-based or aggregated input data, cannot be assessed due to unknown input data accuracies. We could also show that WI differences can only insufficiently be explained by D^{WS} and D^{PS}. Instead, we could show that the spatial accuracy metric *kriging standard deviation* is suitable as an important explaining variable for the non-linear relation between raster- and station-based WIs, which underlines the unsystematic character of WI differences.

Against the background of spatio-temporal quality requirements of geodata in agricultural applications (Möller et al. 2013; Grassini et al. 2015; Lokers et al. 2016; Mourtzinis et al. 2017) and the increasing spatio-temporal availability of climate data (Overpeck et al. 2011), we propose that the derivation of WIs should follow standard protocols including guidelines for an (automatic and efficient) data processing as well as for the standardization of error assessment procedures (Lokers et al. 2016). Since WIs can be considered as (often) complex geodata derivates, all steps of geodata processing should be reproducible and the geodata input data and WIs should be characterized by (spatial) accuracy metrics.

In this study, the derivation of Germany-wide phenological raster data sets followed the idea of a standard protocol since each interpolation result is characterized by the spatial accuracy metric *kriging standard deviation*. This metric was used for the accuracy assessment of phenological windows. As shown on the example of test site 1, variations of phases' starting and end points could be determined, which can consequently affect WI results. This clearly demonstrates that existing WI uncertainties are quantifiable. Here, test site 1 is characterized by WI uncertainties of 21 mm in 2007 and 16 mm in 2013 (see Section 3.3). If (spatial) accuracy metrics for the precipitation data set would be available, than the WI uncertainty could be completely described. By having such information, possible sources of inaccuracies in WI design are identifiable. In this study, this would concern, for instance, the Germany-wide interpolation results of daily temperatures, which are based on (only) 500 weather stations (see Section 2.3.2), but represent a key parameter for the PHASE model (Section 2.2.2).

4.3 WI integrity and trustworthiness

The integrity and traceability of WIs and its geodata components "is highly associated with trust and with having confidence that the quality of data is sufficient to serve as evidence base for critical decision-making (Lokers et al. 2016)." In the context of weather index-based insurances for instance, quantified WI uncertainties, as exemplified in our study, can support the assessment of its effect on resulting payouts to farmers. On the other side, (e.g., institutional) WI providers would also benefit from geodata integrity, since confidence intervals could be defined and WI misinterpretation could be avoided (Lokers et al. 2016).

5 Conclusions and outlook

Germany is the second largest wheat producer in the European Union (FAOSTAT 2015). Lüttger and Feike (2018) found that correlations between winter wheat yields on county level and climatic variables like heat and drought indices vary strongly in terms of time and space. Thus, weather indices (WIs) are needed, which are Germany-wide available and comparable as well as meet requirements regarding standardization, data integrity and trustworthiness.

In this study, we introduced an approach to model dynamic and reproducible weather indices (WIs) based on Germany-wide and freely accessible raster data sets of both daily precipitation and beginning phenological phases. The major advantage of this approach is the possibility to design a spatial and dynamic WI of certain accuracy as well as its standardization.

The presented process chain can be considered as blueprint for the calculation of other WIs based on publicly available data. What remains for further studies is to analyze the quality of the applied WI with regard to its explanatory power for yield fluctuations. Against the background of climate change, WIs gain in importance for analyzing extreme weather conditions and the selection of appropriate measures (Lüttger and Feike 2018). Thus, the further development and improvement of WI accuracies is important not only for science and society but also farmers and WI providers.

Germany can be characterized by a high availability of public and trustworthy meteorological raster data in different temporal resolutions. The interpolation of phenological observations to Germany-wide raster data with spatial accuracy metrics opens up the opportunity of deriving trustworthy WIs. In this context, the Federal Research Centre for Cultivated Plants (JKI) is going to provide both Germany-wide raster data of beginning phenological phases as well as corresponding WIs in the near future (https://emra.julius-kuehn.de).

Acknowledgements We are very grateful to two anonymous reviewers who provided valuable advice on how to improve the manuscript. Parts of this study were funded by the German Federal Ministry of Food and Agriculture and managed by the Federal Office for Agriculture and Food (BLE), contract no. 2815707915.

References

Acevedo E, Silva P, Silva H (2002) Wheat growth and physiology. FAO Plant Production and Protection Series. FAO, Rome, Italy

Adeyinka A, Krishnamurti C, Maraseni T, Chantarat S (2016) The viability of weather-index insurance in managing drought risk in rural Australia. Int J Rural Manag 12:125–142

Barnett BJ, Mahul O (2007) Weather index insurance for agriculture and rural areas in lower-income countries. Am J Agric Econ 89:1241–1247

Castañeda-Vera A, Barrios L, Garrido A, Mínguez I (2014) Assessment of insurance coverage and claims in rainfall related risks in processing tomato in Western Spain. Eur J Agron 59:39–48

Chen W, Hohl R, Tiong L (2017) Rainfall index insurance for corn farmers in shandong based on high-resolution weather and yield data. Agric Finance Rev 77:337–354

Chmielewski FM, Müller A, Bruns E (2004) Climate changes and trends in phenology of fruit trees and field crops in Germany, 1961 - 2000. Agric For Meteorol 121:69–78

Chuine I, Kramer K, Hänninen H (2003) Plant development models Schwartz M (ed), vol 39, An integrative environmental science, Kluwer Academic Publishers, Dordrecht, The Netherlands, Tasks for vegetation science, Phenology

Conradt S, Finger R, Bokuševa R (2015) Tailored to the extremes: Quantile regression for index-based insurance contract design. Agric Econ 46(4):537–547

Conradt S, Finger R, Spörri M (2015) Flexible weather index-based insurance design. Clim Risk Manag 10:106–117

Cruz S, Monteiro A, Santos R (2012) Automated geospatial web services composition based on geodata quality requirements. Comput Geosci 47:60–74

Dalhaus T, Finger R (2016) Can gridded precipitation data and phenological observations reduce basis risk of weather index-based insurance? Weather Clim Soc 8:409–419

Dalhaus T, Musshoff O, Finger R (2018) Phenology Information Contributes to Reduce Temporal Basis Risk in Agricultural Weather Index Insurance. Scientific Reports 8(1):46+. https://doi.org/10.1038/s41598-017-18656-5

Davis J (2002) Statistics and data analysis in geology. John Wiley & Sons

Doms J, Gerstmann H, Möller M (2017) Modeling of dynamic weather indexes by coupling spatial phenological and precipitation data – A practical application in the context of weather index-based insurances. In: Contribution presented at the XV EAAE Congress "Towards Sustainable Agri-food Systems: Balancing Between Markets and Society", European Association of Agricultural Economists (EAAE), Parma, Italy

Doms J, Hirschauer N, Marz M, Boettcher F (2018) Is the hedging efficiency of weather index insurance overrated? A farm-level analysis in regions with moderate natural conditions in Germany. Agric Finance Rev https://doi.org/10.1108/AFR-07-2017-0059

FAOSTAT (2015) FAOSTAT: FAO Statistical database. Tech. rep., Food and Agriculture Organization of the United Nations, Rome, Italy

Field C, Barros V, Stocker T (2012) Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press

Gerstmann H, Doktor D, Gläßer C, Möller M (2016) Phase: A geostatistical model for the kriging-based spatial prediction of crop phenology using public phenological and climatological observations. Comput Electron Agric 127:726–738

Gömann H, Bender A, Bolte A, Dirksmeyer W, Englert H, Feil JH, Frühauf C, Hauschild M, Krengel S, Lilienthal H, Löpmeier FJ, Müller J, Mußhoff O, Natkhin M, Offermann F, Seidel P, Schmidt M, Seintsch B, Steidl J, Strohm K, Zimmer Y (2015) Agrarrelevante Extremwetterlagen und Möglichkeiten von Risikomanagementsystemen: Studie im Auftrag des Bundesministeriums für Ernährung und Landwirtschaft (BMEL), Thünen Rep, vol 30. Johann Heinrich von Thünen-Institut, Braunschweig, Germany

Goodwin B, Mahul O (2004) Risk modeling concepts relating to the design and rating of agricultural insurance contracts. World Bank Policy Research Working Paper 3392, World Bank, Washington, D.C

Grassini P, van Bussel L, Wart JV, Wolf J, Claessens L, Yang H, Boogaard H, de Groot H, van Ittersum M, Cassman K (2015) How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis. Field Crops Res 177:49–63

Hengl T, Heuvelink G, Rossiter D (2007) About regression-kriging: from equations to case studies. Comp Geosci 33(10):1301–1315

Hiemstra P, Pebesma E, Twenhöfel C, Heuvelink G (2009) Real-time automatic interpolation of ambient gamma dose rates from the Dutch Radioactivity Monitoring Network. Comp Geosci 35:1711–1721

Hijmans RJ (2016) raster: Geographic Data Analysis and Modeling. https://CRAN.R-project.org/package=raster, R package version 2.5-8

Kaspar F, Zimmermann K, Polte-Rudolf C (2014) An overview of the phenological observation network and the phenological database of Germany's national meteorological service (Deutscher Wetterdienst). Adv Sci Res 11:93–99

Kuhn M (2008) Building predictive models in R using the caret package. J Stat Softw 28:1–26

Kuhn M, Johnson K (2013) Applied predictive modeling. Springer, New York, Heidelberg, Dordrecht, London

Kuhn M, Wing J, Weston S, Williams A, Keefer C, Engelhardt A, Cooper T, Mayer Z (2014) caret: Classification and Regression Training. http://CRAN.R-project.org/package=caret, R package version 6.0-24

Leblois A, Quirion P (2013) Agricultural insurances based on meteorological indices: realizations, methods and research challenges. Meteorol Appl 20:1–9

Lee JS (1980) Digital image enhancement and noise filtering by use of local statistics. IEEE Trans Pattern Anal Mach Intell 2:165–168

Lokers R, Knapen R, Janssen S, van Randen Y, Jansen J (2016) Analysis of Big Data technologies for use in agro-environmental science. Environ Modell Software 84:494–504

Lüttger AB, Feike T (2018) Development of heat and drought related extreme weather events and their effect on winter wheat yields in Germany. Theor Appl Climatol 132:15–29

McMaster G, Wilhelm W (2003) Phenological responses of wheat and barley to water and temperature: improving simulation models. J Agric Sci 141:129–147

Möller M, Birger J, Gidudu A, Gläßer C (2013) A framework for the geometric accuracy assessment of classified objects. Int J Remote Sens 34:8685–8698

Möller M, Gerstmann H, Gao F, Dahms TC, Förster M (2017) Coupling of phenological information and simulated vegetation index time series: Limitations and potentials for the assessment and monitoring of soil erosion risk. CATENA 150:192–205

Mourtzinis S, Edreira J, Conley S, Grassini P (2017) From grid to field: Assessing quality of gridded weather data for agricultural applications. Eur J Agron 82:163–172

Nellis M, Price K, Rundquist D (2009) Remote sensing of cropland agriculture. In: Warner T, Nellis M, Foody G (eds) The SAGE Handbook of Remote Sensing, vol 1. SAGE Publications, London, UK, pp 368-380

Okpara J, Afiesimama E, Anuforom A, Owino A, Ogunjobi K (2017) The applicability of standardized precipitation index: drought characterization for early warning system and weather index insurance in West Africa. Nat Hazards 89:555–583

Overpeck J, Meehl G, Bony S, Easterling D (2011) Climate data challenges in the 21st century. Sci 331(6018):700–702

Pelka N, Musshoff O (2013) Hedging effectiveness of weather derivatives in arable farming - is there a need for mixed indices? Agric Finance Rev 73:358–372

Pietola K, Myyrä S, Jauhiainen L, Peltonen-Sainio P (2011) Predicting the yield of spring wheat by weather indices in Finland: Implications for designing weather index insurances. Agric Food Sci 20:269–286

Poudel M, Chen S, Huang W (2016) Pricing of rainfall index insurance for rice and wheat in Nepal. J Agric Sci Technol 18:291–302

R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/

Rabus B, Eineder M, Roth A, Bamler R (2003) The shuttle radar topography mission – A new class of digital elevation models acquired by spaceborne radar. ISPRS J Photogramm Remote Sens 57:241–262

Rauthe M, Steiner H, Riediger U, Mazurkiewicz A, Gratzki A (2013) A central European precipitation climatology – Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS). Meteorol Z 22(3):235–256

Rezaei E, Siebert S, Hüging H, Ewert F (2018) Climate change effect on wheat phenology depends on cultivar change. Sci Rep 8(4891)

Schwartz M (ed.) (2006) Phenology: an integrative environmental science, Tasks for Vegetation Science, vol 39, Kluwer Academic Publishers, Dordrecht, The Netherlands

Skees J, Gober S, Varangis P, Lester R, Kalavakonda V (2001) Developing rainfall-based index insurance in Morocco. World Bank Policy Research Working Paper 2577, World Bank, Washington, D.C.

Ssymank A (1994) Neue Anforderungen im europäischen Naturschutz: Das Schutzgebietssystem Natura 2000 und die FFH-Richtlinie der EU. Natur Land 69:395–406

Stoppa A, Hess U (2003) Design and use of weather derivatives in agricultural policies: the case of rainfall index insurance in Morocco. In: International Conference Agricultural Policy Reform and the WTO, Where are we heading, Capri (Italy)

Szoecs E (2016) esmisc: Misc Functions. https://github.com/EDiLD/esmisc/blob/master/R/read_regnie.R, R package version 0.0.2

Turvey CG (2001) Weather derivatives for specific event risks in agriculture. Rev Agric Econ 23:333–351

Vijaya Kumar P, Rao V, Bhavani O, Dubey A, Singh C, Venkateswarlu B (2016) Sensitive growth stages and temperature thresholds in wheat (Triticum aestivum L.) for index-based crop insurance in the Indo-Gangetic Plains of India. J Agric Sci 154:321–333

World Bank (2011) Weather index insurance for agriculture: guidance for development practitioners. No. 50 in Agriculture and Rural Development Discussion Paper World Bank, Washington, D.C

Zhang J, Zhang Z, Tao F (2017) Performance of temperature-related weather index for agricultural insurance of three main crops in China. Int J Disaster Risk Sci 8:78–90

Zhao M, Peng C, Xiang W, Deng X, Tian D, Zhou X, Yu G, He H, Zhao Z (2013) Plant phenological modeling and its application in global climate change research: overview and future challenges. Environ Rev 21:1–14