Data-driven determination of plant growth stages for improved weather index insurance design

Jing Zou
Chair of Statistics and Econometrics, "Friedrich List" Faculty of Transportation, Technische Universität Dresden, Dresden, Germany and Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Dresden/Leipzig, Germany
Martin Odening
Department of Agricultural Economics, Farm Management Group, Humboldt-Universität zu Berlin, Berlin, Germany, and
Ostap Okhrin
Chair of Statistics and Econometrics, "Friedrich List" Faculty of Transportation, Technische Universität Dresden, Dresden, Germany and Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Dresden/Leipzig, Germany

Received 25 January 2024
Revised 30 May 2024
Accepted 17 June 2024

Abstract

Purpose – This paper aims to improve the delimitation of plant growth stages in the context of weather index insurance design. We propose a data-driven phase division that minimizes estimation errors in the weather-yield relationship and investigate whether it can substitute an expert-based determination of plant growth phases. We combine this procedure with various statistical and machine learning estimation methods and compare their performance.

Design/methodology/approach – Using the example of winter barley, we divide the complete growth cycle into four sub-phases based on phenology reports and expert instructions and evaluate all combinations of start and end points of the various growth stages by their estimation errors of the respective yield models. Some of the most commonly used statistical and machine learning methods are employed to model the weather-yield relationship with each selected method we applied.

Findings – Our results confirm that the fit of crop-yield models can be improved by disaggregation of the vegetation period. Moreover, we find that the data-driven approach leads to similar division points as the expert-based approach. Regarding the statistical model, in terms of yield model prediction accuracy, Support Vector Machine ranks first and Polynomial Regression last; however, the performance across different methods exhibits only minor differences.

Originality/value – This research addresses the challenge of separating plant growth stages when phenology information is unavailable. Moreover, it evaluates the performance of statistical and machine learning methods in the context of crop yield prediction. The suggested phase-division in conjunction with advanced statistical methods offers promising avenues for improving weather index insurance design.

Keywords Weather index insurance, Temporal basis risk, Plant growth stages, Generalized additive model, Machine learning

Paper type Research paper

This research is supported by the China Scholarship Council and Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI). We thank the editor and two anonymous reviewers for their comments and suggestions that help improve the quality of this paper.

Agricultural Finance Review
Vol. 84 No. 4/5, 2024
pp. 297-319
Emerald Publishing Limited
0002-1466
DOI 10.1108/AFR-01-2024-0015

1. Introduction

Basis risk, namely the difference between the triggered insurance payout and the actual yield losses, poses a significant hurdle to the wide adoption of weather index insurance (e.g. Woodard and Garcia, 2008; Odening and Shen, 2014; Janzen *et al.*, 2021). Thus, much attention has been paid to quantifying and alleviating this challenge (see Norton *et al.*, 2012; Keller and Saitone, 2022). Basis risk can be decomposed into spatial and product risks (see Boyd *et al.*, 2019).

Spatial basis risk arises from variations between on-site farmland weather conditions and weather indices measured at a reference weather station. These differences can be mitigated through the application of spatial interpolation methods (e.g. Cao *et al.*, 2015; Boyd *et al.*, 2019; Leppert *et al.*, 2021) or the diversification of insurance contracts across different regions (Ritter *et al.*, 2014).

Product basis risk, on the other hand, stems from the inability of weather indices to comprehensively account for variability in agricultural yields. Product basis risk reflects the fact that weather conditions are pivotal in crop production, but are not the sole determinant. Other factors, such as pests, controllable inputs (e.g. fertilizer), and managerial practices affect crop yields as well. Albers *et al.* (2017), for example, estimated that only 43% of the variability in wheat yield can be attributed to weather conditions. Despite this inherent limitation in reducing product basis risk, many attempts have been made to optimize the design of weather index insurance. Two main strands of literature can be distinguished. First, the specification and estimation of the relationship among yield and weather indices. Second, the selection and specification of appropriate weather indices.

The first strand of literature targets reducing product basis risk by estimating the weather-yield relationship with sophisticated statistical modeling. This task is challenging since the functional form of weather-yield relation is not only region- and crop-specific, but, in general, highly nonlinear (e.g. Schlenker and Roberts, 2009). In early works, Polynomial Regression (PR) has been used to model the nonlinear weather-yield relation (e.g. Vedenov and Barnett, 2004). PR is relatively simple in implementation and easy to interpret, however, the resulting model fit is often poor (see Schmidt *et al.*, 2022). In addition, traditional nonparametric regression methods, such as the Generalized Additive Model (GAM) based on B-spline or penalized B-spline (P-spline) smoothers, have also been employed for modeling the nonlinear relationship between weather and yield with a very good quality (cf. Bucheli *et al.*, 2022; Tan and Zhang, 2023; Zou *et al.*, 2023). In addition to conventional statistical methods, more recently, machine learning methods have been applied to weather-yield modeling (Trawiński *et al.*, 2012). In quantifying the effects of weather on crop yield, Random Forest (RF), a tree-based ensemble method (Breiman, 2001), turns out to achieve higher prediction capability than PR (e.g. Feng *et al.*, 2018; Hoffman *et al.*, 2020). Support Vector Machines (SVM) are also regarded as a powerful regression method in crop yield prediction (e.g. Oguntunde *et al.*, 2018; Dang *et al.*, 2021), though they are primarily applied in field crop classification tasks (e.g. Mathur and Foody, 2008; Löw *et al.*, 2013). Moreover, Artificial Neural Networks (ANN) have been used to capture the relationship between high-dimensional weather variables and yield in large data sets (Schmidt *et al.*, 2022; Chen *et al.*, 2023). So far, there is no consensus on the "best" approach for modeling the weather-yield relation and complexity in methods does not always lead to superior performance (Oikonomidis *et al.*, 2022). Some model comparisons have been conducted (e.g. Das *et al.*, 2020; Joshi *et al.*, 2021), but there is no exhaustive comparison of model performance between statistical and machine learning techniques. In addition, machine learning methods are often criticized by their black box properties that imply lack of interpretability, which could probably hinder the promotion of index insurance products among policy holders since contract transparency is often required (Sibiko *et al.*, 2018). To overcome this drawback, visualization with partial dependence plots is considered to increase the explanatory capabilities of machine learning

methods (Greenwell, 2017), which further enables the comparison between methods in terms of the interpretation of results.

The second strand of literature about mitigating product basis risk involves choosing generic weather variables and designing weather indices. There is a consensus that precipitation and temperature are pivotal in crop production. Weather indices are usually calculated by aggregating weather variables (e.g. cumulative rainfall or average temperature). Often used weather indices are Growing Degree Days (GDD), Cumulative Rainfall (CR), and the standardized precipitation index (cf. Turvey, 2001; Stoppa and Hess, 2003; Hill *et al*., 2019; Okpara *et al*., 2017). Moreover, remote sensing data are used to retrieve vegetation indices as alternatives to meteorological indices in the design of index insurance (cf. Möllmann *et al*., 2019; Vroege *et al*., 2021). Important design parameters are the beginning and length of the accumulation period and the aggregation level of weather indices. Turvey *et al*. (2021) select the start and end months of a crop growing season from all combination candidates based on model performance. If the accumulation period of the weather indices is not aligned with the vegetation period of the insured crop, a so-called temporal basis risk may arise (Conradt *et al*., 2015a), which can be considered a special case of product basis risk. Moreover, the varying water demand of crops and vulnerability against weather events during different plant growth stages are often ignored (Deng *et al*., 2007). To address these problems and thus mitigate temporal basis risk, it has been suggested to split the entire vegetation period into distinct plant growth stages and to calculate weather indices for each growth stage separately (e.g. Zou *et al*., 2023). Usually, the determination of beginning and end dates of growth phases is conducted by experts based on phenology information. However, the required information is not always available, at least not at the desired spatial resolution. This deficit calls for a data-driven division procedure as an alternative.

Against this backdrop, the contribution of our paper is twofold. First, we propose a data-driven method to search for the optimal division points of plant growth stages as an alternative to an expert-based growth stage division. This is accomplished via the minimization of weather-yield relation estimation errors. We investigate the feasibility as well as the costs and benefits of this approach. We find that the benefits are modest in terms of improvements in model fit. However, model development cost and computational costs are outweighed in situations where external information on division timings does not exist. Second, we provide a thorough comparison of classical statistical methods with various machine learning techniques in the context of yield prediction with weather variables. The comparison not only focuses on model fit and estimation errors, but also includes the visualization of the partial dependence of crop yields on individual weather variables in different growth phases. We use relatively simple, commonly used weather indices because our study focuses on a comparison of statistical methods and the optimal decomposition of the vegetation period.

The remainder of this paper is organized as follows: Section 2 motivates the selected statistical and machine learning techniques with a simulation study illustrating the strengths and weaknesses of each method; Section 3 displays the data and study region of our empirical analysis; Section 4 describes model estimation including parameter tuning procedures, shows numerical results, and visualizes and interprets the impact of individual predictors on response variable by means of partial dependence plots; and Section 5 concludes.

2. Statistical methods and machine learning techniques for weather-yield modeling

2.1 Method selection

In this study, we implement five methods for modeling crop weather-yield relation: Polynomial Regression (PR), Generalized Additive Model (GAM), Random Forest (RF), Support Vector Machines (SVM), and Artificial Neural Networks (ANN). The model choice is

based on the idea that the methods should represent different model families, specifically statistical methods and machine learning techniques, and should also complement each other. Moreover, the selection considers model performance in previous analyses. PR is a parametric linear model and serves as the complementary benchmark in this study for the evaluation of the machine learning methods. A penalization technique, namely the Least Absolute Shrinkage and Selection Operator (LASSO) (Tibshirani, 1996), is excluded from the regression since feature selection by reducing insignificant redundant variables does not apply to our case given at most eight input weather indices. Note that although PR is used to estimate non-linear relationships with quadratic terms, it is still a linear regression given the linearity in regression coefficients (Hastie *et al.*, 2009a). GAM is a family member of generalized linear models (Hastie and Tibshirani, 1985), which possesses the interpretability of linear models but in a non-parametric setting (Fan and Jiang, 2005) and exhibits good performance (Zou *et al.*, 2023). In our study, the smoothers in GAM are constructed by two-dimensional P-splines, which are considered as a state-of-art statistical method in yield modeling (cf. Tan and Zhang, 2023). The other three models, RF, SVM and ANN, are classical tools in machine learning and have been proven to be able to model general non-linear relationships. RF is a representative of tree-based methods and, turns out to be a robust prediction method compared to many other tree-based models. For example, Extreme Gradient Boosting (XGBoost) proposed by Chen and Guestrin (2016) is prone to overfitting when the sample size of training data is small as pointed out by Folberth *et al.* (2019). Considering SVM's excellent capability in training a small-sized data set (Mountrakis *et al.*, 2011) and that yield data are usually scarce (Odening and Shen, 2014), SVM is selected from kernel-based learning methods to solve the regression task in this study. Finally, ANN offers a comparative machine learning technique to RF and SVM, although it is more widely implemented in solving big data problems (e.g. Pasini, 2015). In our application, a pair of indices $WI1$ and $WI2$ are calculated simultaneously in each index aggregation period, such as the accumulative precipitation and average temperature in a certain month, during one plant growth stage or across the whole growth cycle. Since this study highlights the phase-division procedure, we do not dive into the selection of weather indices, which can be investigated in further research. We estimate the expected values of yield variables although the estimation of the lower bound of the yield distribution might be useful in an insurance context (e.g. Conradt *et al.*, 2015b). Moreover, in a decision theoretic framework, e.g. expected utility maximization, information about the entire yield distribution is required (e.g. Zhang *et al.*, 2019). Likewise, the decomposition of basis risk into upside and downside basis risk hinges on the entire yield distribution (cf. Schmidt *et al.*, 2022). In what follows, we describe the specification of the five selected methods in greater detail.

2.1.1 Polynomial regression (PR). PR is a widely used statistical model that represents a linear combination of predictors, their powers, and cross-products, offering the advantages of simplicity and interpretability. For example, Vedenov and Barnett (2004) apply PR to model the nonlinear relationships between weather indices and corn, cotton, and soybean yields. Chen *et al.* (2017) design rainfall index insurance using PR. In this paper, we employ quadratic regression to estimate nonlinear relations as a benchmark following Schmidt *et al.* (2022), where PR fails to capture the yield variability as substantially as the machine learning methods. To illustrate this, we assume only one index aggregation period. For example, in case $WI1$ and $WI2$ are two different weather indices calculated across the whole growth cycle, the PR model takes the following form:

$$Y_{m,l} = \beta + \beta_1 WI1_{m,l} + \beta_2 WI2_{m,l} + \beta_3 WI1_{m,l}^2 + \beta_4 WI2_{m,l}^2 + \beta_5 WI1_{m,l} \cdot WI2_{m,l} + \epsilon_{m,l}, \quad (1)$$

where $Y_{m,l}$ is yield-related response variable, m represents the year, l denotes the farm, β is the intercept, and $\beta_1, \ldots, \beta_5$, are coefficients estimated by Ordinary Least Squares. Note that when

the number of phases is more than one, the cross-products are measured within each aggregation period, not across crop development stages.

2.1.2 Generalized additive model (GAM). GAM is a regression model encompassing a linear combination of smoothers (Hastie and Tibshirani, 1987). This paper utilizes Penalized B-spline (P-spline) tensor products as smoothers since they are advantageous in modeling nonparametric interactions (cf. Lee *et al.*, 2013). For instance, Tan and Zhang (2023) model the interplay between precipitation and vapor pressure deficits on yield losses using a B-spline tensor product and penalty function. Zou *et al.* (2023) implement a P-spline tensor product to design index insurance with GDD and CR. Despite the intuitive visualization, the model performance was not satisfactory in their case. Again, we use a B-spline tensor product as the type of smoother, $S(WI1_{m,l}, WI2_{m,l})$, under the whole growth cycle case. To construct a B-spline tensor product, we first conduct the de-Boor recursion (de Boor, 1978) to calculate the marginal B-spline basis $N_{j,d}^{u}$ of each input weather covariate as follows:

$$N_{j,0}(u) = \begin{cases} 1, if\ u_j \leq u < u_{j+1} \\ 0, otherwise \end{cases} \text{ and} \tag{2}$$

$$N_{j,d} = \frac{u - u_j}{u_{j+d} - u_j} N_{j,d-1}(u) + \frac{u_{j+d+1} - u}{u_{j+d+1} - u_{j+1}} N_{j+1,d-1}(u), \tag{3}$$

where u represents the element in the weather index vector, the order $d = 1, 2, 3$ when selecting cubic spline, $\{u_j\}$, $j = 0, \ldots, nk - 1$, is a uniform knot vector generated by the equidistant knots in the range of the weather index vector, and nk is the number of knots (cf. Tan and Zhang, 2023). Secondly, a P-spline tensor product is defined as the Kronecker product of B-spline marginal bases along two dimensions of weather indices with penalization. Being consistent with the variables and denotations in 2.1.1, a GAM that contains one smoother is $Y_{m,l} = S(WI1_{m,l}, WI2_{m,l}) + \varepsilon'_{m,l}$.

2.1.3 Random forest (RF). As introduced by Breiman (2001), the RF algorithm first draws a bootstrap sample B times from the dataset. This involves randomly selecting observations with replacements to create each bootstrap sample. These bootstrap samples are then used to grow individual trees in the RF ensemble $T(WI_b; \Theta_b), b = 1, \ldots, B$, where: WI_b represent the b th group of input weather indices; Θ_b is the b th independent and identically distributed random vector constituted of corresponding split variables, cut points at each node, and terminal values of the tree; and B is the total number of trees (Hastie *et al.*, 2009b). Finally, predictions are obtained by averaging the ensemble of unweighted trees, i.e. $Y_{m,l} = \frac{1}{B} \sum_{b=1}^{B} T(WI_{m,l,b}; \Theta_{m,l,b})$. Through this process, the algorithm builds RF and identifies the best-split variable and cut points to divide nodes into two daughter ones, iteratively continuing until reaching the minimum node size. Prasad *et al.* (2021) found that the performance of RF is reliable and faster in predicting cotton yield. Schierhorn *et al.* (2021) also implement RF to explore the contribution of weather to winter wheat yields across various crop growth stages. Nevertheless, RF shows poor extrapolation capabilities when the training data are narrowly distributed (Montes *et al.*, 2021).

2.1.4 Support vector machines (SVM). While initially introduced for classification tasks (Cortes and Vapnik, 1995), SVM has gained popularity in regression applications (Gunn, 1998). By leveraging SVM, data points are projected into a higher-dimensional space using kernel techniques (Meyer, 2023). To better illustrate the application of SVM in weather-yield regression, we take the whole-cycle (non-phase-division) case for denotation simplicity. First, we reprocess the input weather indices $WI1_{m,l}$ and $WI2_{m,l}$, by mapping them into a feature space (Aizerman, 1964) and the map is denoted as $\Phi(WI1_{m,l}, WI2_{m,l})$. The SVM algorithm depends on the dot

product of this map, which is defined as $Y_{m,l} = \langle \Phi(WI1_{m,l}, WI2_{m,l}), \Phi(WI1_{m,l}^T, WI2_{m,l}^T) \rangle$, where the superscript T denotes the transpose of the covariate vector. However, it could be computationally infeasible to derive the explicit forms of both the map function and the dot product, especially in the case of complex nonlinearity (Smola and Schölkopf, 2004). To solve this problem, "kernel tricks" are implemented to replace the dot product with a kernel function (Brereton and Lloyd, 2010), which is represented as $K(WI, WI^T) = \langle \Phi(WI), \Phi(WI^T) \rangle$, where WI denotes the matrix of weather covariates $[WI1_{m,l}, WI2_{m,l}]$. Therefore, the optimization task can be accomplished with this kernel function and a commonly utilized one is the radial basis function (cf. Dang *et al.*, 2021; Lischeid *et al.*, 2022). For details on the optimization process, refer to Smola and Schölkopf (2004). Some studies conclude that SVM regression can achieve more accurate results in agricultural yield prediction (e.g. Brdar *et al.*, 2011; Shah *et al.*, 2018). Chen *et al.* (2016) also find that SVM performs dominantly in selecting climate variables that impact yield variability. SVM handles small training data sets effectively, but the estimation results are highly susceptible to hyper-parameter settings (Mountrakis *et al.*, 2011).

2.1.5 Artificial neural networks (ANN). ANNs are designed to simulate the functioning of neurons in the human brain within a computer system (Krogh, 2008). The network architecture typically consists of the number of layers (such as input, hidden, and output layers), the number of neurons within each layer, the activation function used in each layer, and the training algorithm employed (Benardos and Vosniakos, 2007). The construction component, namely the artificial neuron, is defined as $Y_{m,l} = g\left(\sum_{h=1}^{H} w_{m,l,h} WI_{m,l,h} + bias\right)$, where $g(\cdot)$ is the activation function, h denotes the ordinal number of the input weather index, H is the total number of input variables, w_h is the weight coefficient associated with the h th weather index WI_h, and *bias* denotes the biases. Schmidt *et al.* (2022) employed the rectifier linear unit (ReLU) activation function, i.e. $g(v) = \max(v, 0)$ (Sharma *et al.*, 2017), to estimate the nonlinear weather-yield relationship. Moreover, the Adam optimizer (Kingma and Ba, 2014) was utilized in the training process. By leveraging ANN, intricate and nontrivial patterns in the data can be captured, allowing for a more comprehensive and accurate modeling of the weather-yield relationship (Schmidt *et al.*, 2022; Chen *et al.*, 2023). However, ANNs are commonly used in big data analysis and often more factors need to be taken into consideration when implementing ANNs in small sample sizes (e.g. Pasini, 2015; Shaikhina and Khovanova, 2017).

2.2 Simulation study

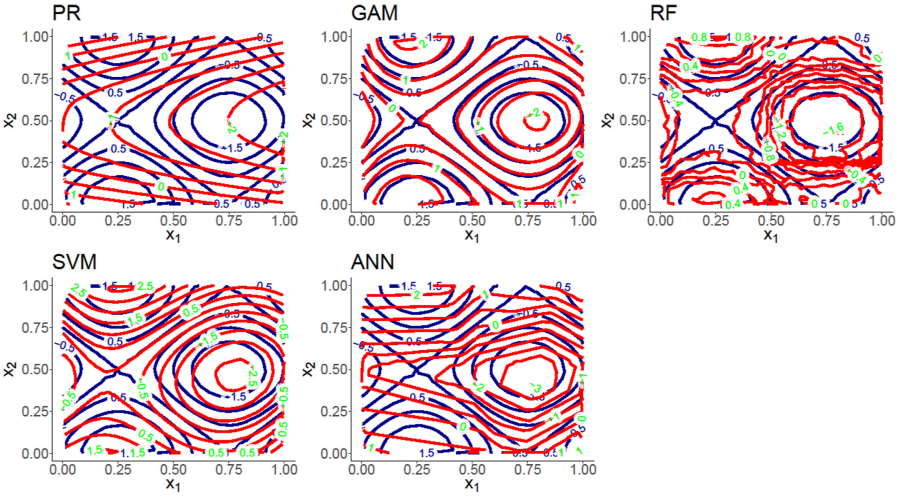
Despite the wide use of the above methods in weather-yield modeling, no model has been recognized as superior. It calls for methodological guidance through a thorough comparison between these methods, which is one of the contributions in this study. To illustrate their characteristics and potential flaws, we conduct a simulation study with artificial data generated by four nonlinear functions (the functional form is specified in the Appendix [cf. Lee *et al.*, 2013]). We sampled uniformly distributed inputs $x_i, i = 1, \ldots, 8$ with a sample size of 1,000 and evaluated the performance with the added white noises. The value of the function filled with the sampled observation is considered as a response variable. Note that for the hyper-parameter tuning process of SVM and ANN, identical training and test datasets were employed. The results are in Table A1 in the Appendix. The simulation study results are presented in Figure 1, which showcases the contour plots of true functions in blue and the estimated surfaces produced by the five methods in red fixing $x_3, x_4, x_5, x_6, x_7, x_8 = 0.50$. GAM demonstrates its robustness in approximating the true surface, while PR, RF, SVM, and ANN exhibit noticeable discrepancies consistent with the descriptions of methods above. First, PR shows the most prominent performance deficiency in predicting nonlinear relations.

Second, analyzing the limited sample size, SVM outperforms ANN. Third, as shown by the vacant area around the point ($x_1 = 0.25$, $x_2 = 0.50$) RF is insufficient in extrapolation. These findings highlight the varying performance of the different methods in capturing the underlying relationships between covariates, which motivate the objective of comparing methods in this empirical study. All the models have problems in performing well at the boundaries due to lacking degrees of freedom.

This simulation study was intended to show weak points of the methods. In real-world applications, the function linking weather and yield are unknown. Therefore, we need to rely on other benchmarks or comparisons, keeping in mind the models' weaknesses.

3. Data and study region

Our study area, the state of Saxony, is located in eastern Germany. According to the Saxon State Office for the Environment, Agriculture and Geology, a total of 12 agro-ecological zones are defined across the state based on soil properties and climate conditions. Figure 2 illustrates the distribution of these agro-ecological zones in polygon form. From north to south, soil is separated into three regions with relatively fertile soils in the central part. For details on the soil conditions, please refer to the map in the Appendix. According to the 2020 agricultural census in Germany, about 6,500 farms in Saxony cultivate around 894,100 hectares of arable land. The most important grains are wheat with around 182,700 hectares and barley with 115,200 hectares. The most important oil crop, winter rapeseed, is grown on 101,500 hectares (StLa Sachsen, 2020). From an email survey conducted in Saxony in 2022 by the Saxon State Ministry for Energy, Climate Protection, Environment and Agriculture, weather risks, e.g. hail, drought, frost, and ice days, are highly agriculturally relevant with drought ranked as the most common agricultural risk that resulted in yield losses. Nonetheless, the surveyed farmers were rather dissatisfied with the market's risk management instruments. At the same time, they rated insurance as an attractive instrument, especially in view of the increase in weather risks caused by climate change. This real-world demand motivates our research to design a weather index insurance product for crop producers in Saxony.



Source(s): Figure created by authors

Figure 1. Contour plots of $\widehat{f}(x_1, x_2, x_3 = x_4 = x_5 = x_6 = x_7 = x_8 = 0.5)$ estimated by Polynomial Regression (PR), Generalized Additive Model (GAM), Random Forest (RF), Support Vector Machines (SVM), and Artificial Neural Networks (ANN), respectively

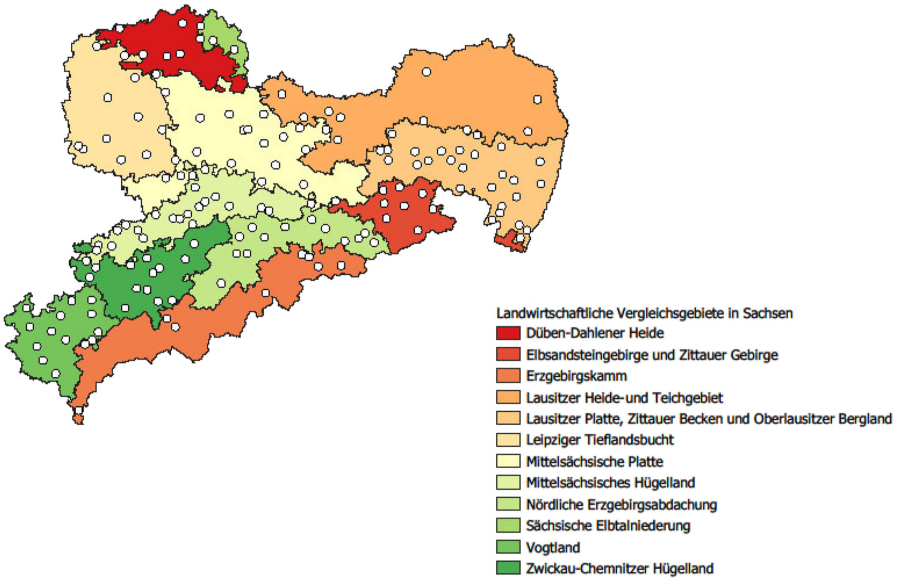


Figure 2.
Agro-ecological zone classification in the state of Saxony

Note(s): Each color denotes an agro-ecological zone (Landwirtschaftliches Vergleichsgebiet), and the white points represent 157 municipalities

Source(s): Figure created by authors

Our empirical analysis incorporates yield, meteorological, and phenological data. Yield data were obtained from the annual financial statements of the Federal Ministry of Food and Agriculture through the Farm Accountancy Data Network, which is a microeconomic data source that relies on national surveys within Europe. Farm-level yield data provide advantages over county-level aggregations when estimating yield variability, as highlighted in previous studies (Popp *et al.*, 2005; Schmidt *et al.*, 2022). In our case study, we focus on winter barley [1] for two reasons. First, our data set is most complete for this crop, which are measured at the farm level and originate from 217 farms located in 157 municipalities across 12 agro-ecological zones in Saxony. The data spans from 2000 to 2015. For confidentiality reasons, the exact coordinates of farm locations are concealed, but information about the municipality a farm is located in is available. In Figure 2, these municipalities are shown as white points. Second, winter barley had a significant cultivation in Germany between 1989 and 2020 (Ellsäßer, 2022). Moreover, winter barley is a globally important cereal crop, widely used in animal feed, beer, and spirits production (Verstegen *et al.*, 2014). Thus, there is potential demand for weather index insurance that protects winter barley producers in Saxony against weather perils. For a robustness check of our findings, we estimate the weather-yield relation for rapeseed, which is also an important economic crop in Saxony. The rapeseed yield data in our data set cover 202 farms located in 142 municipalities from 2000 to 2015 and the results are presented in the Appendix.

In this study we define yield deviations $Y_{m,l} = DY_{m,l} - \overline{Y}_l$ as the response variable, where m represents the year, l denotes the farm, $\overline{Y}_l$ is the average yield of farm l, $DY_{m,l}$ are detrended yield observations, and negative deviations indicate farm losses. We detrended the raw yield observations $YO_{m,l}$ using a linear time regression model $YO_{m,l} = \alpha + \delta(m - 2000) + DY_{m,l}$ (Shi and Jiang, 2016). The detailed descriptive statistics of $YO_{m,l}$, $DY_{m,l}$, and $Y_{m,l}$ are displayed in Table 1.

Daily weather data for Germany, including precipitation, maximum and minimum temperatures, are publicly available through the Deutscher Wetterdienst (DWD). Specifically, in the daily Hydrometeorologische Rasterdaten (HYRAS) data sets of the DWD that we use, the precipitation product for Germany is in a 1 × 1 km grid and the spatial resolution of the temperature raster is 5 × 5 km. To unify the spatial resolution of raw weather files, we use resampling techniques to obtain the 1 × 1 km gridded temperature data. To relate the weather data with municipalities, we calculate the average values in the polygons of municipalities and export the means.

In this study, we choose Growing Degree Days (GDD) and Cumulative Rainfall (CR) as weather indices. GDD and CR are calculated according to the following definitions:

$$GDD_{m,l,s} = \sum_{t=sd_s}^{ed_s} max\left\{0, \frac{(TMA_{m,l,s,t} + TMI_{m,l,s,t})}{2} - TB\right\} \text{ and} \tag{4}$$

$$CR_{m,l,s} = \sum_{t=sd_s}^{ed_s} R_{m,l,s,t}, \tag{5}$$

where $s \in \{1, 2, 3, 4\}$ denotes the ordinal number of winter barley growth stage, t is the t-th day in a year, sd and ed denote the start and the end dates, respectively, of the accumulation period, whose unit is day of the year (DOY), and TMA, TMI, and R are the daily maximum temperature, minimum temperature, and rainfall, respectively. The baseline temperature TB of winter barley is defined as 0° Celsius (Manderscheid *et al.*, 2009). Start dates of the accumulation periods are not dependent on the year m and we assume no temporal overlapping between two consecutive phases, i.e. $sd_{s+1} = ed_s$ for $s \in \{1, 2, 3\}$. To calculate the weather indices GDD and CR for each farm, we average the gridded daily weather data collected from the DWD HYRAS data sets in each municipality instead of each farm because exact farm locations are unknown. However, the geographical basis risk is negligible due to the relatively small size of the municipality.

DWD also provides phenology reports in Saxony, but the spatial resolution of the data is not satisfactory. Although there are data from 57 phenology observation stations available, the number of available phenological stations varies each year with phenology timing. When we link each municipality with the nearest phenology station to determine the beginning and end dates, a discontinuity of data would be introduced. Thus, for each phase in each year, we define the medians of observations from available stations across Saxony and ignore spatial heterogeneity. Likewise, the temporal resolution of the phenology information is not expedient because it is not available on a weekly basis (Vallentin *et al.*, 2022). To overcome this limitation, we follow the procedures in Schmitt *et al.* (2022) by deriving dates based on historical time distances between consecutive phenological phases and expert knowledge.

	Yield (dt/ha)	Detrended yield (dt/ha)	Yield deviations (dt/ha)
Mean	61.76	55.24	0
Standard deviation	17.03	12.74	12.74
Minimum[a]	0	−4.37	−59.60
Maximum	122.88	122.63	67.39

Note(s): [a]Given that we use farm yield data and the fact that farms in East Germany encountered severe droughts in several years within the observation period, a zero yield observation is not unlikely. It simply means that the yield was so low that it was not profitable to incur harvesting costs

Source(s): Table created by authors

Table 1. Descriptive statistics of yield data

Specifically, we first obtain the medians of all observed data as state-level phenology dates as described above. Then, we add a certain number of days to estimate the missing phenological information. For instance, the timing of "flowering" sd_2 is defined by experts as the observation data of "heading" plus 11 days, while the occurrence of "fruit formation" sd_3 is considered as 21 days after "heading". Therefore, the temporal distance between "flowering" and "fruit formation" is as long as ten days across 16 years. The determination of phase length is highly dependent on expert experience. Thus, it can introduce additional errors due to individual cognition biases. The lack of appropriate phenology data is exactly the reason that motivates a data-driven growth phase determination.

To test the effectiveness of the data-driven phase division, we specify three distinct models:

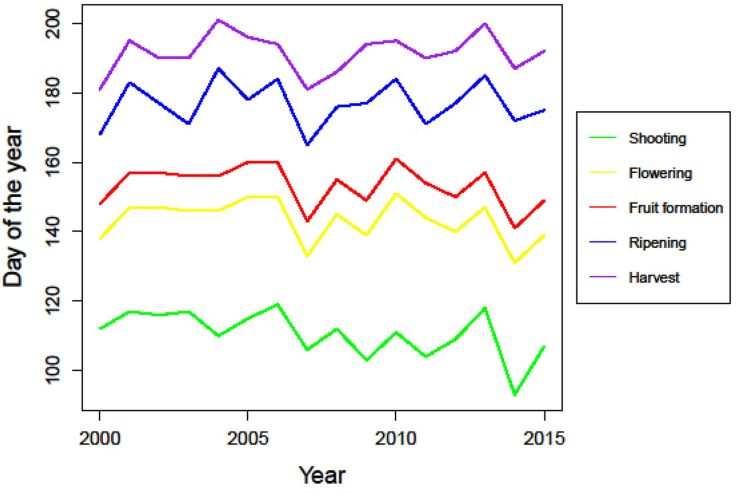
(1) A model without separation of growth phase ("non-division", hereafter "ND").

(2) A model with exogenously determined length of growth stages ("expert-division", hereafter "ED").

(3) A model with data-driven determination of growth phase length ("optimal-division", hereafter "OD").

The first two models serve as benchmarks for the third. Models (2) and (3) are referred to as "phase-division" models.

Following the methodology proposed by Schmitt *et al.* (2022), we divide the entire growth cycle of winter barley into four stages and the intervals of stage 1 (from "shooting" to "flowering"), stage 2 (from "flowering" to "fruit formation"), stage 3 (from "fruit formation" to "ripening"), and stage 4 (from "ripening" to "harvest") are $[sd_1, ed_1]$, $[sd_2, ed_2]$, $[sd_3, ed_3]$, and $[sd_4, ed_4]$, respectively. Figure 3 shows DOYs for each growth phase from 2000 to 2015. Throughout the 16 years, the earliest occurrence of the "shooting" stage was on the 93rd DOY in 2014, while the latest occurrence of the "harvest" stage was on the 201st DOY in 2004. Therefore, we define the start and end division points as $sd_1 = 93$ and $ed_4 = 201$, respectively. Moreover, we denote the median DOY of internal division dates across 16 years as $\widetilde{sd}_2 = 145$, $\widetilde{sd}_3 = 155$, and $\widetilde{sd}_4 = 177$. sd_1 and ed_4 mark the beginning and end of the



Source(s): Figure created by authors

Figure 3.
Day of the year (DOY) of phenological phenomenon timing

vegetation period in the non-division model. Note that these two timestamps are constant and identical in all models. In the expert-division model, the three internal division points are fixed at $\widetilde{sd}_2, \widetilde{sd}_3$, and $\widetilde{sd}_4$. With these specifications, GDD and CR time series are calculated for all ND and ED models. Table 2 displays the descriptive statistics of these weather indices. It indicates that the non-division weather indices vary greatly from expert-division ones regarding mean, standard deviation, minimal value, and maximal value. In addition, significant difference among stages in the statistics of expert-division weather indices is shown (see Table 2).

To conduct an optimal phase division, we parametrize sd_2, sd_3, and sd_4 and enlarge the search space to k possible combinations, which amount to 24,804 combinations of growth phase durations in this study, given the constraints of minimal phase lengths, namely:

$$\begin{cases} sd_2 \geq sd_1 + ml_1 - 1 \\ sd_3 \geq sd_2 + ml_2 - 1 \\ sd_4 \geq sd_3 + ml_3 - 1 \\ sd_4 \leq ed_4 - ml_4 + 1 \end{cases}, \tag{6}$$

where ml_s, $s = 1, 2, 3, 4$ denotes the minimal phase length of the s-th stage and ml_1, ml_2, ml_3, and ml_4 are 26, 10, 15, and 10, respectively (see Figure 3). For each method, the optimal combination of sd_2, sd_3, and sd_4 is denoted as $(\ddot{sd}_2, \ddot{sd}_3, \ddot{sd}_4)$, respectively. This combination is selected by maximizing adjusted R^2 or minimizing the Root Mean Squared Error (RMSE), i.e. $RMSE = \sqrt{\sum_{m=1}^{n}\sum_{l=1}^{217}\left(\frac{Y_{m,l} - \widehat{Y}_{m,l})}{217n}\right)^2}$, where n is the length of time series in each farm in the training data set, m denotes the year, l represents the farm, $Y_{m,l}$ are the yield deviations, and $\widehat{Y}_{m,l}$ are the fitted values.

4. Model estimation and results

4.1 Parameter tuning

As described in Section 3, three types of models, namely non-division (ND), expert-division (ED), and optimal-division (OD) models, are estimated in this study. Among the five methods implemented, PR, GAM, and RF can be used without parameter tuning. All computations were performed using the *R* statistical software (version 4.3.1). For GAM, we use the gam() function using the P-spline and REML criterion in the "mgcv" package. RFs were estimated using the "randomForest" package (Breiman, 2001). The number of trees is set to 500 and the number of variables at each split is 2 (cf. Biau and Scornet, 2016). In contrast, the estimations of the two machine learning methods, SVM and ANN, contain an extra process of hyper-parameter selection.

Table 2. Descriptive statistics of time series of growing degree days (GDD) (in degree Celsius) and cumulative rainfall (CR) (in mm) for non-division (ND) and expert-division (ED) models

	ND-GDD	ED-GDD_1	ED-GDD_2	ED-GDD_3	ED-GDD_4	ND-CR	ED-CR_1	ED-CR_2	ED-CR_3	ED-CR_4
Mean	1509.06	575.88	155.44	376.70	446.70	234.86	83.86	34.92	53.01	68.74
Standard deviation	105.11	65.44	28.35	39.85	42.73	66.75	28.17	41.31	26.63	35.68
Minimum	1120.83	371.50	78.79	255.64	293.29	71.86	24.43	0.00	0.69	2.21
Maximum	1746.02	749.97	217.01	459.30	557.07	547.86	187.32	250.04	159.80	195.98

Source(s): Table created by authors

As introduced in Section 3, for each type of model, we calculate accumulative measures GDD and CR and the input data are normalized into the range $[0, 1]$ for computational stability. With constant division points, the weather indices of ND and ED models are static and the hyper-parameter tuning procedures required by SVM and ANN are accomplished with a normal cross validation process.

However, for the optimal-division model, we obtain k combinations of both the division points and corresponding weather indices, which implies k times of cross validation and estimation for each method before we determine the optimal-division points in terms of adjusted R^2 or *RMSE* as introduced in Section 3. For the methods that require no parameter tuning, namely PR, GAM, and RF, it is not an overly complex task, but for SVM and ANN, it would be too time consuming, since $k = 24,804$ in our case. To balance the trade-off between costs and hyper-parameter accuracy, we first selected 20 division points combinations randomly from k possibilities. Second, for each selected division points combination, we selected hyper-parameters via cross-validation. We divided the dataset into a training set spanning 11 years and a test set of 5 years, which ensures a substantial test dataset size of 1,085 (the product of 5 years and 217 farms), that avoids misevaluation, e.g. spurious higher accuracy in machine learning implementation (Vabalas *et al.*, 2019). Afterward, results of 20 sets of hyper-parameters were averaged for evaluating k models.

In the cross-validation process, we selected a radial basis kernel function for SVM based on the methodology described in Lischeid *et al.* (2022). The parameters "cost" and "gamma" were determined through grid search. Regarding ANN, we designed a fully connected ANN with two hidden layers and determined the hyper-parameters using grid search employing the ReLU activation function (Sharma *et al.*, 2017) and the Adam optimizer (Kingma and Ba, 2014). The number of neurons were selected from 4, 8, and 16 in two-dimensional ND models. In the two phase-division models, the options were 8, 16, and 32. We also incorporated ridge regression as a penalty term corresponding to L_2 regularization to prevent overfitting. The results of the parameter tuning process in empirical study are provided in Table A2 in the Appendix.

4.2 Model results

After the estimations with ND, ED, and OD (24,804 rounds) models using each method, we obtained out-of-sample results (displayed in Table 3). Note that in the ND model estimated by ANN, we used the sigmoid activation function instead of ReLU because our parameter tuning tests indicate that the model performance is unstable using ReLU in a two-dimensional case. Moreover, the ND model was not estimated by RF since with two independent variables, the RF algorithm would randomly select at most one predictor from the two predictors as a subset and thus cannot further select the best-split variable recursively.

Method	ND	*RMSE*	Adj. R^2	ED	*RMSE*	Adj. R^2	OD	*RMSE*	Adj. R^2
PR		12.00	0.11		10.76	0.28	(139, 149, 182)	10.32	0.34
GAM		11.75	0.15		10.00	0.38	(139, 149, 177)	9.61	0.43
RF	–	–	–	(145, 155, 177)	9.80	0.41	(121, 132, 180)	9.74	0.41
SVM		11.77	0.14		9.01	0.50	(132, 141, 189)	7.62	0.64
ANN		11.75	0.15		9.11	0.49	(151, 160, 177)	9.51	0.44

Table 3. $(\ddot{sd}_2, \ddot{sd}_3, \ddot{sd}_4)$, *RMSE*, and adjusted R^2 of ND, ED, and OD models

Note(s): ND, ED, and OD are abbreviations for non-division, expert-division, and optimal-division. PR, GAM, RF, SVM, and ANN are abbreviations for Polynomial Regression, Generalized Additive Model, Random Forest, Support Vector Machines, and Artificial Neural Networks, respectively
Source(s): Table created by authors

We compare the model performance and the division points based on the results. First, Table 3 shows that *RMSE* is around 10 dt/ha in most cases. This indicates a high basis risk from weather index insurance since 10 dt/ha is about 18% of the average detrended yield for winter barley (55.24 dt/ha). This reflects the general limitation of the prediction capability of weather indices for farm level yields. For the ND models, the performance differences of the four methods are negligible. In both phase-division models, the results differ more across methods, but they are still minor. Overall, PR has a relatively weak model fit, which is in line with earlier studies (e.g. Shah *et al.*, 2018). In contrast, SVM ranks first among the five methods using separate growth phases. This confirms the finding of a review study of SVM in agricultural applications (Kok *et al.*, 2021). GAM and RF are rather similar regarding model performance.

Table 3 further shows that phase-division models significantly outperform models that do not disaggregate the vegetation period into growth phases. The relative reduction in *RMSE* ranges from 14.0% to 35.2%, confirming the advantage of the phenology-based disaggregation in the weather index calculation of the estimation of the weather-yield relation that has been reported in previous studies, albeit with different study areas and crops (Zou *et al.*, 2023; Dalhaus *et al.*, 2018). Comparing the two phase-division models shows that model fit can be further improved by the proposed data-driven phase division, at least in combination with PR, GAM, and SVM. The performance gain form relaxing the internal three points is most significant for SVM. In contrast, for ANN this procedure comes along with a performance loss compared to the expert-based growth division model. Overall, the gains of optimal phase division compared to ED are only modest. However, our results demonstrate the potential of the data-driven phase-division as an alternative to expert-based phase division.

Apart from the model performance, it is interesting to compare the phase-division points derived by the different methods with the expert-based division points, i.e. $\widetilde{sd}_2 = 145$, $\widetilde{sd}_3 = 155$, and $\widetilde{sd}_4 = 177$ .We find that the combinations generated by PR [139, 149, 182], GAM [139, 149, 177], and ANN [151, 160, 177] deviate the least from the ED growth phases. Specifically, the $\ddot{sd}_4$ of GAM and ANN is exact to the median $\widetilde{sd}_4 = 177$. Another remarkable finding is that under OD, $\ddot{sd}_4$ of PR, GAM, RF, SVM, and ANN are 182, 177, 180, 189, and 177, respectively. This suggests that significant consistency exists in the third optimal point $\ddot{sd}_4$, whereas it does not in the other two. However, since this study aims to search for the data-driven division points determined by the model selection, the optimal points chosen by each method do not need to be consistent with the expert-division ones.

Table 4 displays the lengths of the growth phases that result from the two approaches. The five statistical methods suggest different lengths, i.e. there is no unique OD. Only for the second phase do all methods suggest a similar length. This can be explained by the minimum

Division model	Stage 1	Stage 2	Stage 3	Stage 4
ED	52	10	22	24
OD-PR	47	11	34	20
OD-GAM	47	11	29	25
OD-RF	29	12	49	22
OD-SVM	40	10	49	13
OD-ANN	59	10	18	25

Note(s): ED and OD are abbreviations for expert-division and optimal-division. PR, GAM, RF, SVM, and ANN are abbreviations for Polynomial Regression, Generalized Additive Model, Random Forest, Support Vector Machines, and Artificial Neural Networks, respectively
Source(s): Table created by authors

Table 4.
ED and OD phase lengths (in days)

length setting of 10 days, as introduced in Section 3. Due to data scarcity of phenology information from "flowering" to "fruit formation", the period is fixed to be 10 days long across 16 years according to the expert definition. Thus, when defining the search space of $(\ddot{sd}_2, \ddot{sd}_3, \ddot{sd}_4)$, one constraint is that $ml_2 = 10$ in Inequality (6). The results of the second phase in Table 4 indicate convergence to a short duration for winter barley for blooming and beginning the grain filling. For stage 1 (from "shooting" to "flowering"), the data driven determination results in a shorter length (except ANN) than the ED phase length, while for stage 3 (from "fruit formation" to "ripening") the opposite is true. The regularity revealed by the data-driven results as opposed to expert knowledge can be explained by global warming induced by climate change. Increasing temperatures accelerate the flowering process (Distefano *et al.*, 2018), slow the fruit formation period since it is more water intensive, and likely harm the soil's water storage capacity (Werner *et al.*, 2020).

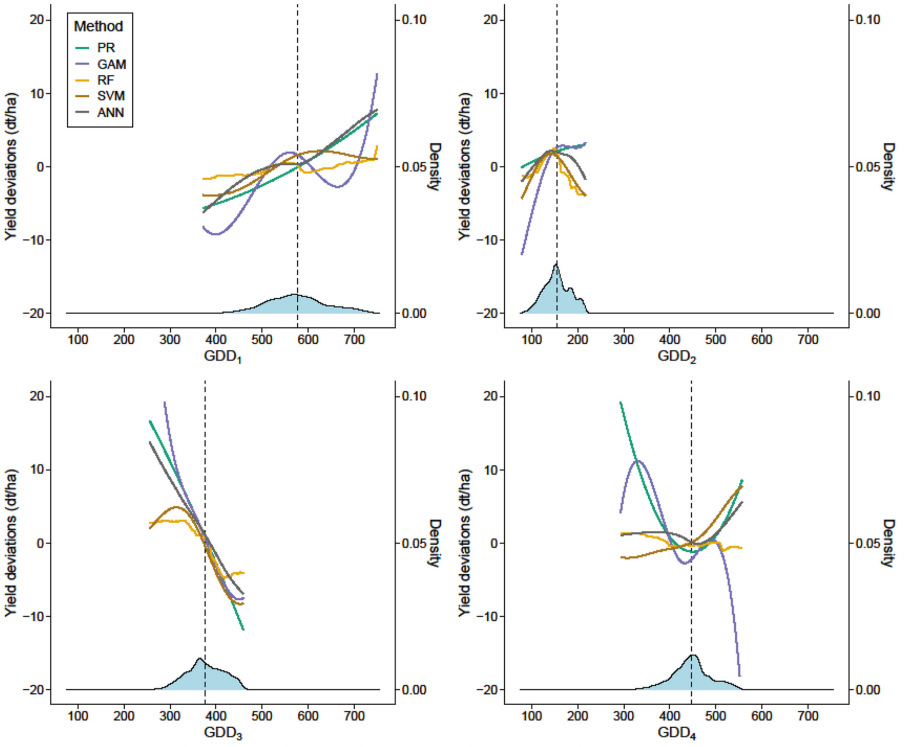
4.3 Model interpretation

Machine learning methods are often considered as "black boxes" with a non-transparent link between input and output variables (e.g. Castelvecchi, 2016). Farmers, however, are interested in understanding the impact of weather events on crop yields, not only to verify indemnity payments from index insurance, but also to support managerial decisions, such as about irrigation and fertilization. For this purpose, we display partial dependence plots that predict the marginal effect of each explanatory variable on the response variable, while controlling for the other predictors through averaging their effects (Friedman, 2001). In the context of crop yield modeling, partial dependence plots have been employed by Jeong *et al.* (2016) and Beillouin *et al.* (2020).

Figures 4 and 5 show the partial dependence plot of the two weather indices (GDD and CR) for all growth stages. The *y*-axis on the left represents the yield deviations corresponding to the weather index value on the *x*-axis, excluding the effects of other predictors. The *y*-axis on the right side denotes the density of the weather index.

Figure 4 visualizes the effect of temperature on yield deviations. It is striking that the impact of GDD differs between growth stages. In phase 1 (from "shooting" to "flowering"), higher GDD_1 are associated with greater yields, while the opposite is true at the third stage (from "fruit formation" to "ripening"). These statistical findings can be explained from an agronomic perspective. In the early growth stage, high temperatures facilitate blooming and development after the vernalization process (Amasino and Michaels, 2010). In the second growth phase (from "flowering" to "fruit formation"), partial dependence plots of GDD_2 for PR and GAM are rather flat when the index value exceeds its mean. This is consistent with the results reported in Schlenker and Roberts (2009) where the yield growth curves of US corn, soybeans, and cotton all exhibit relatively flat shapes. In contrast, in our study, the plots of RF, SVM, and ANN show an inverse U-shape, meaning that deviations from an average temperature level in both directions cause yield losses. These losses, however, are rather small within the observed range of the GDD. The negative effect of higher temperature in phase 3 is consistent with the finding from the data-driven phase length in Section 4.2, i.e. decreasing soil water holding ability can accompany heat stress (Wahid *et al.*, 2007; Werner *et al.*, 2020). At the fourth stage (from "ripening" to "harvest"), the relatively flat partial dependence plots of ANN, RF, and SVM imply that impacts of GDD on yield deviations are modest when winter barley matures. The irregular and U-shaped plots produced by GAM and PR appear implausible.

Figure 5 depicts the partial dependence of winter barley yield on cumulative rainfall in the respective growth phases. For the first growth stage, all plots exhibit a negative slope as long as CR_1 is below its mean value. This confirms the well-known fact that insufficient precipitation in the early plant growth phase causes yield losses (e.g. Stephens and Lyons,



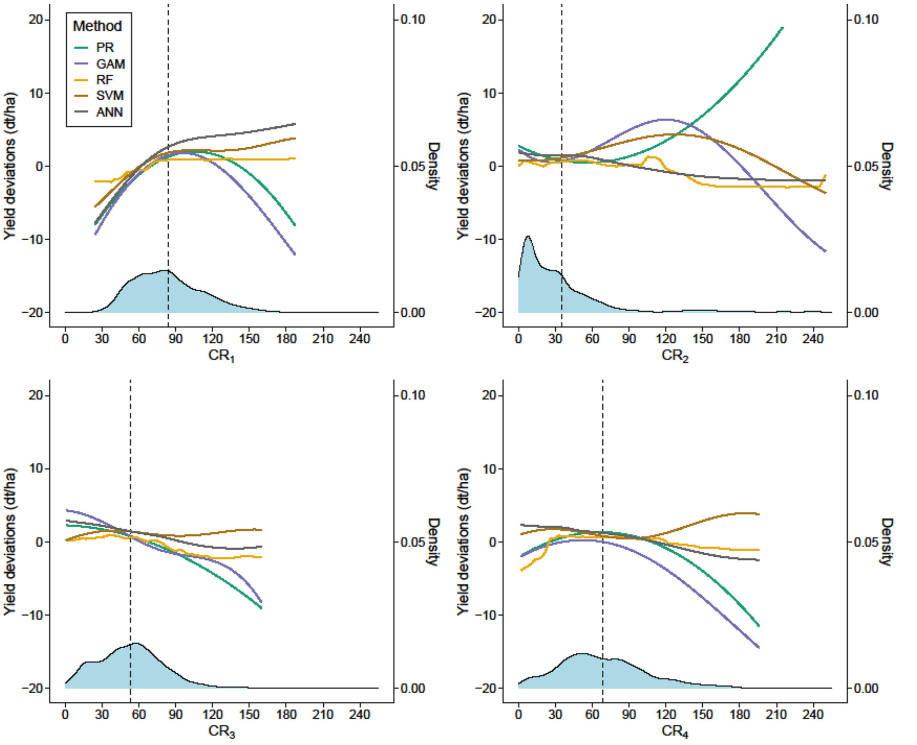
Note(s): The notation under the X-axis refers to the weather index of each phase, e.g., GDD_1 corresponds to phase 1 (from "shooting" to "flowering"). The dashed line indicates the mean value of GDD at each stage. PR, GAM, RF, SVM, and ANN are abbreviations for Polynomial Regression, Generalized Additive Model, Random Forest, Support Vector Machines, and Artificial Neural Networks, respectively

Source(s): Figure created by authors

Figure 4.
Partial dependence plots of growing degree days (GDD) in four growth stages

1998). The curves of ANN, RF, and SVM are rather flat around the mean values in stages 2, 3, and 4, indicating robustness of winter barley yields against drought as well as excessive rainfall. This "crop yield plateau effect" has also been reported in Turvey *et al.* (2021). One explanation is that the length of the second growth phase is relatively short (10 days) compared with the first one (52 days). One should also recall that these plots do not account for interactions among growth stages. For example, due to the water storage capacity of soil, rainfall in the first period can be carried over to subsequent periods. Using soil moisture instead of cumulative rainfall might capture these temporal spillover effects. Moreover, we cannot rule out that some farms use irrigation to compensate for rainfall deficits.

Overall, the partial dependence plots show similar patterns across the five statistical methods, indicating the robustness of the results – at least in the neighborhood of the indices' mean values. However, some abnormal deviations occur for PR and GAM in some growth stages (e.g. GDD_4 and CR_2) if the indices take extreme values. PR behaves wildly because few observations on the boundaries have very little effect on the Ordinary Least Squares, but the shape of the function is determined as quadratic polynomial. Specifically, the curve of GAM shows more variation than the other four methods. In Alimadad and Salibian-Barrera (2011),



Note(s): The notation under the X-axis refers to the weather index of each phase, e.g., CR_1 corresponds to phase 1 (from "shooting" to "flowering"). The dashed line indicates the mean value of CR at each stage. PR, GAM, RF, SVM, and ANN are abbreviations for Polynomial Regression, Generalized Additive Model, Random Forest, Support Vector Machines, and Artificial Neural Networks, respectively

Source(s): Figure created by authors

Figure 5.
Partial dependence plots of cumulative rainfall (CR) in each growth stage

GAM can be very sensitive to outliers, in spite of their small proportion in the observations. The instability of partial dependence plots can be considered as a reflection of the limitation of these methods. Moreover, some methods result in dependence plots that are almost constant in the tails of the distribution, i.e. RF for $GDD_1 < 400$ in Figure 4. The RF curve stays constant if there is a lack of data because the fitted values are obtained from average trees that provide constant values at the leaves. These extrapolation failures of tree-based methods, such as RF, together with other data-driven methods, such as SVM and ANN, have been proved in previous studies (cf. Hettiarachchi *et al.*, 2005; Hengl *et al.*, 2018).

5. Conclusions

This research was motivated by the observation that the performance of crop yield models can be improved by a decomposition of the entire vegetation period into distinct growth phase in which weather events may have a different effect on crop yields. This, in turn, facilitates the design of effective weather index insurance. Unfortunately, the required phenological information for the determination of the beginning and end of each growth stage

is not readily available. This study proposes a novel data-driven growth phase division that does not hinge on expert-based information.

Our study offers several findings relevant for crop yield modelers, weather insurance developers, and farmers. First, we confirm the expectation that a breakdown of the vegetation period and the disaggregation of weather indices results in a better model fit. Second, the endogenization and optimization of division points leads only to a slight improvement compared with an expert-based determination. This arouses a discussion on the cost-benefit ratio of the optimal division procedure. Regarding the costs of the suggested data-driven determination of growth phases, one has to distinguish between computational costs, which are rather low, and development costs, which may be considerable. These costs have to be compared with the labor costs of an *in situ* data collection of phenological information. The benefits, on the other hand, depend on the availability of an expert-based phase division. If insurance providers have access to phenological information, there is only little incentive to develop a model with endogenous splitting dates for the growth phases, because the gain in model performance is minor. However, if expert knowledge is not available, which is often the case in developing countries, the data-driven phase division constitutes a feasible alternative. It is also noteworthy that the optimal division points generated by the various statistical methods differ, i.e. there is no unique optimal division. Nonetheless, model performance is rather robust against changes in the timing and length of the phases. Third, regarding the statistical approach, we find that the parametric PR model is outperformed by nonparametric and non-linear ones, i.e. either spline-based GAM or machine learning models (RF, ANN, SVM). In our application, SVM show the best model fit though they have primarily been designed for classifications tasks. Despite the initial constraints of our study, which focused on a small geographical area and a limited time interval, our findings indicate that the application of smooth, flexible methods, namely GAM, RF, SVM, and ANN, yielded comparable results in terms of R^2 and RMSE. Notably, we acknowledge that we did not explicitly test for statistical significance. Although ANN posed challenges in terms of training and tuning due to its computational intensity, the similarity in partial dependence plots among RF, SVM, and ANN suggests a comparable performance across these models. It would be interesting to examine whether these findings can be generalized to other crops and other climatic regions.

Despite of the aforementioned improvements, one has to acknowledge that the basis risk of a weather insurance product derived from these models is still high. Even in the best case, average prediction errors amount to more than 10% of the average barley yield and undermine the hedging effectiveness of index insurance. In fact, the uptake of index-based weather insurance in the study region is modest and higher subsidies from the government would be necessary to increase the adoption of this kind of index insurance product in reality (Odening *et al.*, 2022). However, there are several opportunities for optimization to further reduce basis risk. First, the inclusion of additional or more sophisticated weather variables (e.g. rainfall deficit, dry spells, soil moisture, excessive rainfall, and heat days) will most likely enhance model performance (e.g. Bucheli *et al.*, 2021). Second, satellite data have received considerable research attention to reduce the geographical basis risk (e.g. Möllmann *et al.*, 2019; Vroege *et al.*, 2021). Third, as pointed out by Bucheli *et al.* (2021), considerable risk reduction can be achieved if the underlying drought index is tailored individually for each farm. Finally, the parameter tuning process of SVM and ANN could be further refined.

Notes

1. Winter wheat is the most important crop in Saxony and we have considered analyzing this crop, but unfortunately in our data set winter wheat is merged with summer wheat. Estimating a yield model for this aggregate would be inappropriate due to the different vegetation periods of winter and summer wheat.

References

Aizerman, A. (1964), "Theoretical foundations of the potential function method in pattern recognition learning", *Automation and Remote Control*, Vol. 25, pp. 821-837.

Albers, H., Gornott, C. and Hüttel, S. (2017), "How do inputs and weather drive wheat yield volatility? The example of Germany", *Food Policy*, Vol. 70, pp. 50-61, doi: 10.1016/j.foodpol.2017.05.001.

Alimadad, A. and Salibian-Barrera, M. (2011), "An outlier-robust fit for generalized additive models with applications to disease outbreak detection", *Journal of the American Statistical Association*, Vol. 106 No. 494, pp. 719-731, doi: 10.1198/jasa.2011.tm09654.

Amasino, R.M. and Michaels, S.D. (2010), "The timing of flowering", *Plant Physiology*, Vol. 154 No. 2, pp. 516-520, doi: 10.1104/pp.110.161653.

Beillouin, D., Schauberger, B., Bastos, A., Ciais, P. and Makowski, D. (2020), "Impact of extreme weather conditions on European crop production in 2018", *Philosophical Transactions of the Royal Society B*, Vol. 375 No. 1810, 20190510, doi: 10.1098/rstb.2019.0510.

Benardos, P.G. and Vosniakos, G.C. (2007), "Optimizing feedforward artificial neural network architecture", *Engineering Applications of Artificial Intelligence*, Vol. 20 No. 3, pp. 365-382, doi: 10.1016/j.engappai.2006.06.005.

Biau, G. and Scornet, E. (2016), "A random forest guided tour", *Test*, Vol. 25 No. 2, pp. 197-227, doi: 10.1007/s11749-016-0481-7.

Boyd, M., Porth, B., Porth, L. and Turenne, D. (2019), "The impact of spatial interpolation techniques on spatial basis risk for weather insurance: an application to forage crops", *North American Actuarial Journal*, Vol. 23 No. 3, pp. 412-433, doi: 10.1080/10920277.2019.1566074.

Brdar, S., Culibrk, D., Marinkovic, B., Crnobarac, J. and Crnojevic, V. (2011), "Support vector machines with features contribution analysis for agricultural yield prediction", in *The Second International Workshop on Sensing Technologies in Agriculture*, Forestry and Environment (EcoSense 2011), Belgrade, pp. 43-47.

Breiman, L. (2001), "Random forests", *Machine Learning*, Vol. 45 No. 1, pp. 5-32, doi: 10.1023/a:1010933404324.

Brereton, R.G. and Lloyd, G.R. (2010), "Support vector machines for classification and regression", *Analyst*, Vol. 135 No. 2, pp. 230-267, doi: 10.1039/b918972f.

Bucheli, J., Dalhaus, T. and Finger, R. (2021), "The optimal drought index for designing weather index insurance", *European Review of Agricultural Economics*, Vol. 48 No. 3, pp. 573-597, doi: 10.1093/erae/jbaa014.

Bucheli, J., Dalhaus, T. and Finger, R. (2022), "Temperature effects on crop yields in heat index insurance", *Food Policy*, Vol. 107, 102214, doi: 10.1016/j.foodpol.2021.102214.

Cao, X., Okhrin, O., Odening, M. and Ritter, M. (2015), "Modelling spatio-temporal variability of temperature", *Computational Statistics*, Vol. 30 No. 3, pp. 745-766, doi: 10.1007/s00180-015-0561-2.

Castelvecchi, D. (2016), "Can we open the black box of AI?", *Nature News*, Vol. 538 No. 7623, pp. 20-23, doi: 10.1038/538020a.

Chen, T. and Guestrin, C. (2016), "Xgboost: a scalable tree boosting system", *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785-794.

Chen, H., Wu, W. and Liu, H.B. (2016), "Assessing the relative importance of climate variables to rice yield variation using support vector machines", *Theoretical and Applied Climatology*, Vol. 126 Nos 1-2, pp. 105-111, doi: 10.1007/s00704-015-1559-y.

Chen, W., Hohl, R. and Tiong, L.K. (2017), "Rainfall index insurance for corn farmers in Shandong based on high-resolution weather and yield data", *Agricultural Finance Review*, Vol. 77 No. 2, pp. 337-354, doi: 10.1108/afr-10-2015-0042.

Chen, Z., Lu, Y., Zhang, J. and Zhu, W. (2023), "Managing weather risk with a neural network-based index insurance", *Management Science*, Vol. 70 No. 7, pp. 4306-4327, doi: 10.1287/mnsc.2023.4902.

Conradt, S., Finger, R. and Spörri, M. (2015a), "Flexible weather index-based insurance design", *Climate Risk Management*, Vol. 10, pp. 106-117, doi: 10.1016/j.crm.2015.06.003.

Conradt, S., Finger, R. and Bokusheva, R. (2015b), "Tailored to the extremes: quantile regression for index-based insurance contract design", *Agricultural Economics*, Vol. 46 No. 4, pp. 537-547, doi: 10.1111/agec.12180.

Cortes, C. and Vapnik, V. (1995), "Support-vector networks", *Machine Learning*, Vol. 20 No. 3, pp. 273-297, doi: 10.1007/bf00994018.

Dalhaus, T., Musshoff, O. and Finger, R. (2018), "Phenology information contributes to reduce temporal basis risk in agricultural weather index insurance", *Scientific Reports*, Vol. 8 No. 1, p. 46, doi: 10.1038/s41598-017-18656-5.

Dang, C., Liu, Y., Yue, H., Qian, J. and Zhu, R. (2021), "Autumn crop yield prediction using data-driven approaches:-support vector machines, random forest, and deep neural network methods", *Canadian Journal of Remote Sensing*, Vol. 47 No. 2, pp. 162-181, doi: 10.1080/07038992.2020.1833186.

Das, B., Nair, B., Arunachalam, V., Reddy, K.V., Venkatesh, P., Chakraborty, D. and Desai, S. (2020), "Comparative evaluation of linear and nonlinear weather-based models for coconut yield prediction in the west coast of India", *International Journal of Biometeorology*, Vol. 64 No. 7, pp. 1111-1123, doi: 10.1007/s00484-020-01884-2.

de Boor, C. (1978), *A Practical Guide to Splines*, Springer, New York.

Deng, X., Barnett, B.J., Vedenov, D.V. and West, J.W. (2007), "Hedging dairy production losses using weather-based index insurance", *Agricultural Economics*, Vol. 36 No. 2, pp. 271-280, doi: 10.1111/j.1574-0862.2007.00204.x.

Distefano, G., Gentile, A., Hedhly, A. and La Malfa, S. (2018), "Temperatures during flower bud development affect pollen germination, self-incompatibility reaction and early fruit development of clementine (Citrus clementina Hort. ex Tan.)", *Plant Biology*, Vol. 20 No. 2, pp. 191-198, doi: 10.1111/plb.12656.

Ellsäßer, F. (2022), "Cropdata–spatial yield productivity data base for the ten most cultivated crops in Germany from 1989 to 2020-version 1.0", doi: 10.22029/jlupub-7177.

Fan, J. and Jiang, J. (2005), "Nonparametric inferences for additive models", *Journal of the American Statistical Association*, Vol. 100 No. 471, pp. 890-907, doi: 10.1198/016214504000001439.

Feng, P., Wang, B., Liu, D.L., Xing, H., Ji, F., Macadam, I., Ruan, H. and Yu, Q. (2018), "Impacts of rainfall extremes on wheat yield in semi-arid cropping systems in eastern Australia", *Climatic Change*, Vol. 147 Nos 3-4, pp. 555-569, doi: 10.1007/s10584-018-2170-x.

Folberth, C., Baklanov, A., Balkovič, J., Skalský, R., Khabarov, N. and Obersteiner, M. (2019), "Spatio-temporal downscaling of gridded crop model yield estimates based on machine learning", *Agricultural and Forest Meteorology*, Vol. 264, pp. 1-15, doi: 10.1016/j.agrformet.2018.09.021.

Friedman, J.H. (2001), "Greedy function approximation: a gradient boosting machine", *Annals of Statistics*, Vol. 29 No. 5, pp. 1189-1232, doi: 10.1214/aos/1013203451.

Greenwell, B.M. (2017), "pdp: an R package for constructing partial dependence plots", *R Journal*, Vol. 9 No. 1, p. 421, doi: 10.32614/rj-2017-016.

Gunn, S.R. (1998), "Support vector machines for classification and regression", *ISIS Technical Report*, Vol. 14 No. 1, pp. 5-16.

Hastie, T. and Tibshirani, R. (1985), "Generalized additive models; some applications", *Generalized Linear Models: Proceedings of the GLIM 85 Conference held at Lancaster, UK, Sept. 16-19, 1985*, Springer US, pp. 66-81.

Hastie, T. and Tibshirani, R. (1987), "Generalized additive models: some applications", *Journal of the American Statistical Association*, Vol. 82 No. 398, pp. 371-386, doi: 10.2307/2289439.

Hastie, T., Tibshirani, R. and Friedman, J. (2009a), "Linear methods for regression", *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, pp. 43-99.

Hastie, T., Tibshirani, R. and Friedman, J. (2009b), "Random forests", in *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, pp. 587-604.

Hengl, T., Nussbaum, M., Wright, M.N., Heuvelink, G.B. and Gräler, B. (2018), "Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables", *PeerJ*, Vol. 6, e5518, doi: 10.7717/peerj.5518.

Hettiarachchi, P., Hall, M.J. and Minns, A.W. (2005), "The extrapolation of artificial neural networks for the modelling of rainfall—runoff relationships", *Journal of Hydroinformatics*, Vol. 7 No. 4, pp. 291-296, doi: 10.2166/hydro.2005.0025.

Hill, R.V., Kumar, N., Magnan, N., Makhija, S., de Nicola, F., Spielman, D.J. and Ward, P.S. (2019), "Ex ante and ex post effects of hybrid index insurance in Bangladesh", *Journal of Development Economics*, Vol. 136, pp. 1-17, doi: 10.1016/j.jdeveco.2018.09.003.

Hoffman, A.L., Kemanian, A.R. and Forest, C.E. (2020), "The response of maize, sorghum, and soybean yield to growing-phase climate revealed with machine learning", *Environmental Research Letters*, Vol. 15 No. 9, 094013, doi: 10.1088/1748-9326/ab7b22.

Janzen, S., Magnan, N., Mullally, C., Shin, S., Palmer, I.B., Oduol, J. and Hughes, K. (2021), "Can experiential games and improved risk coverage raise demand for index insurance? Evidence from Kenya", *American Journal of Agricultural Economics*, Vol. 103 No. 1, pp. 338-361, doi: 10.1111/ajae.12124.

Jeong, J.H., Resop, J.P., Mueller, N.D., Fleisher, D.H., Yun, K., Butler, E.E., Timlin, D.J., Shim, K.M., Gerber, J.S., Reddy, V.R. and Kim, S.H. (2016), "Random forests for global and regional crop yield predictions", *PloS One*, Vol. 11 No. 6, e0156571, doi: 10.1371/journal.pone.0156571.

Joshi, V.R., Kazula, M.J., Coulter, J.A., Naeve, S.L. and Garcia y Garcia, A. (2021), "In-season weather data provide reliable yield estimates of maize and soybean in the US central Corn Belt", *International Journal of Biometeorology*, Vol. 65 No. 4, pp. 489-502, doi: 10.1007/s00484-020-02039-z.

Keller, J.B. and Saitone, T.L. (2022), "Basis risk in the pasture, rangeland, and forage insurance program: evidence from California", *American Journal of Agricultural Economics*, Vol. 104 No. 4, pp. 1203-1223, doi: 10.1111/ajae.12282.

Kingma, D.P. and Ba, J. (2014), "Adam: a method for stochastic optimization", arXiv Preprint arXiv: 1412.6980.

Kok, Z.H., Shariff, A.R.M., Alfatni, M.S.M. and Khairunniza-Bejo, S. (2021), "Support vector machine in precision agriculture: a review", *Computers and Electronics in Agriculture*, Vol. 191, 106546, doi: 10.1016/j.compag.2021.106546.

Krogh, A. (2008), "What are artificial neural networks?", *Nature Biotechnology*, Vol. 26 No. 2, pp. 195-197, doi: 10.1038/nbt1386.

Lee, D.J., Durbán, M. and Eilers, P. (2013), "Efficient two-dimensional smoothing with P-spline ANOVA mixed models and nested bases", *Computational Statistics and Data Analysis*, Vol. 61, pp. 22-37, doi: 10.1016/j.csda.2012.11.013.

Leppert, D., Dalhaus, T. and Lagerkvist, C.J. (2021), "Accounting for geographic basis risk in heat index insurance: how spatial interpolation can reduce the cost of risk", *Weather, Climate, and Society*, Vol. 13 No. 2, pp. 273-286, doi: 10.1175/wcas-d-20-0070.1.

Lischeid, G., Webber, H., Sommer, M., Nendel, C. and Ewert, F. (2022), "Machine learning in crop yield modelling: a powerful tool, but no surrogate for science", *Agricultural and Forest Meteorology*, Vol. 312, 108698, doi: 10.1016/j.agrformet.2021.108698.

Löw, F., Michel, U., Dech, S. and Conrad, C. (2013), "Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using support vector machines", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 85, pp. 102-119, doi: 10.1016/j.isprsjprs.2013.08.007.

Manderscheid, R., Pacholski, A., Frühauf, C. and Weigel, H.J. (2009), "Effects of free air carbon dioxide enrichment and nitrogen supply on growth and yield of winter barley cultivated in a crop rotation", *Field Crops Research*, Vol. 110 No. 3, pp. 185-196, doi: 10.1016/j.fcr.2008.08.002.

Mathur, A. and Foody, G.M. (2008), "Crop classification by support vector machine with intelligently selected training data for an operational application", *International Journal of Remote Sensing*, Vol. 29 No. 8, pp. 2227-2240, doi: 10.1080/01431160701395203.

Meyer, D., (2008), "Support vector machines", *The interface to libsvm in package e1071*, available at: https://ftp.gwdg.de/pub/misc/cran/web/packages/e1071/vignettes/svmdoc.pdf (accessed 5 December 2023).

Möllmann, J., Buchholz, M. and Musshoff, O. (2019), "Comparing the hedging effectiveness of weather derivatives based on remotely sensed vegetation health indices and meteorological indices", *Weather, Climate, and Society*, Vol. 11 No. 1, pp. 33-48, doi: 10.1175/wcas-d-17-0127.1.

Montes, C., Kapelan, Z. and Saldarriaga, J. (2021), "Predicting non-deposition sediment transport in sewer pipes using Random forest", *Water Research*, Vol. 189, 116639, doi: 10.1016/j.watres.2020.116639.

Mountrakis, G., Im, J. and Ogole, C. (2011), "Support vector machines in remote sensing: a review", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 66 No. 3, pp. 247-259, doi: 10.1016/j.isprsjprs.2010.11.001.

Norton, M.T., Turvey, C. and Osgood, D. (2012), "Quantifying spatial basis risk for weather index insurance", *The Journal of Risk Finance*, Vol. 14 No. 1, pp. 20-34, doi: 10.1108/15265941311288086.

Odening, M. and Shen, Z. (2014), "Challenges of insuring weather risk in agriculture", *Agricultural Finance Review*, Vol. 74 No. 2, pp. 188-199, doi: 10.1108/afr-11-2013-0039.

Odening, M., Filler, G., Schmidt, L., Musshoff, O., Michels, M., Mahler, S., Böttcher, F. (2022), "Mehrgefahrenversicherung in der Landwirtschaft im Freistaat Sachen. Feasibility Study on Multi-peril crop insurance on behalf of the Saxonian Ministry of Energy, Climate Protection", Environment and Agriculture. Mimeo.

Oguntunde, P.G., Lischeid, G. and Dietrich, O. (2018), "Relationship between rice yield and climate variables in southwest Nigeria using multiple linear regression and support vector machine analysis", *International Journal of Biometeorology*, Vol. 62 No. 3, pp. 459-469, doi: 10.1007/s00484-017-1454-6.

Oikonomidis, A., Catal, C. and Kassahun, A. (2022), "Hybrid deep learning-based models for crop yield prediction", *Applied Artificial Intelligence*, Vol. 36 No. 1, e2031822, doi: 10.1080/08839514.2022.2031823.

Okpara, J.N., Afiesimama, E.A., Anuforom, A.C., Owino, A. and Ogunjobi, K.O. (2017), "The applicability of Standardized Precipitation Index: drought characterization for early warning system and weather index insurance in West Africa", *Natural Hazards*, Vol. 89 No. 2, pp. 555-583, doi: 10.1007/s11069-017-2980-6.

Pasini, A. (2015), "Artificial neural networks for small dataset analysis", *Journal of Thoracic Disease*, Vol. 7 No. 5, pp. 953-960, doi: 10.3978/j.issn.2072-1439.2015.04.61.

Popp, M., Rudstrom, M. and Manning, P. (2005), "Spatial yield risk across region, crop and aggregation method", *Canadian Journal of Agricultural Economics/Revue Canadienne D'agroeconomie*, Vol. 53 Nos 2-3, pp. 103-115, doi: 10.1111/j.1744-7976.2005.00408.x.

Prasad, N.R., Patel, N.R. and Danodia, A. (2021), "Crop yield prediction in cotton for regional level using random forest approach", *Spatial Information Research*, Vol. 29 No. 2, pp. 195-206, doi: 10.1007/s41324-020-00346-6.

Ritter, M., Musshoff, O. and Odening, M. (2014), "Minimizing geographical basis risk of weather derivatives using a multi-site rainfall model", *Computational Economics*, Vol. 44 No. 1, pp. 67-86, doi: 10.1007/s10614-013-9410-y.

Schierhorn, F., Hofmann, M., Gagalyuk, T., Ostapchuk, I. and Müller, D. (2021), "Machine learning reveals complex effects of climatic means and weather extremes on wheat yields during different plant developmental stages", *Climatic Change*, Vol. 169 No. 39, doi: 10.1007/s10584-021-03272-0.

Schlenker, W. and Roberts, M.J. (2009), "Nonlinear temperature effects indicate severe damages to US crop yields under climate change", *Proceedings of the National Academy of sciences*, Vol. 106 37, pp. 15594-15598, doi: 10.1073/pnas.0906865106.

Schmidt, L., Odening, M., Schlanstein, J. and Ritter, M. (2022), "Exploring the weather-yield nexus with artificial neural networks", *Agricultural Systems*, Vol. 196, 103345, doi: 10.1016/j.agsy.2021.103345.

Schmitt, J., Offermann, F., Söder, M., Frühauf, C. and Finger, R. (2022), "Extreme weather events cause significant crop yield losses at the farm level in German agriculture", *Food Policy*, Vol. 112, 102359, doi: 10.1016/j.foodpol.2022.102359.

Shah, A., Dubey, A., Hemnani, V., Gala, D. and Kalbande, D.R. (2018), "Smart farming system: crop yield prediction using regression techniques", *Proceedings of International Conference on Wireless Communication: ICWiCom 2017*, Singapore, Springer, pp. 49-56.

Shaikhina, T. and Khovanova, N.A. (2017), "Handling limited datasets with neural networks in medical applications: a small-data approach", *Artificial Intelligence in Medicine*, Vol. 75, pp. 51-63, doi: 10.1016/j.artmed.2016.12.003.

Sharma, S., Sharma, S. and Athaiya, A. (2017), "Activation functions in neural networks", *International Journal of Engineering Applied Sciences and Technology*, Vol. 4 No. 12, pp. 310-316, doi: 10.33564/ijeast.2020.v04i12.054.

Shi, H. and Jiang, Z. (2016), "The efficiency of composite weather index insurance in hedging rice yield risk: evidence from China", *Agricultural Economics*, Vol. 47 No. 3, pp. 319-328, doi: 10.1111/agec.12232.

Sibiko, K.W., Veettil, P.C. and Qaim, M. (2018), "Small farmers' preferences for weather index insurance: insights from Kenya", *Agriculture and Food Security*, Vol. 7, pp. 1-14, doi: 10.1186/s40066-018-0200-6.

Smola, A.J. and Schölkopf, B. (2004), "A tutorial on support vector regression", *Statistics and Computing*, Vol. 14 No. 3, pp. 199-222, doi: 10.1023/b:stco.0000035301.49549.88.

Stephens, D.J. and Lyons, T.J. (1998), "Rainfall-yield relationships across the Australian wheatbelt", *Australian Journal of Agricultural Research*, Vol. 49 No. 2, pp. 211-224, doi: 10.1071/a96139.

StLa Sachsen (2020), "Presse, Medieninformation 92/2020 vom 29.07.2020, Landwirtschaftszählung 2020: weniger Getreide und Zuckerrüben, mehr Raps und Hülsenfrüchte auf sächsischen Feldern", available at: https://www.statistik.sachsen.de/download/presse-2020/mi_statistik-sachsen_092-2020_landwirtschaftszaehlung-2020.pdf (accessed 22 June 2024).

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

Tan, K.S. and Zhang, J. (2023), "Flexible weather index insurance design with penalized splines", *North American Actuarial Journal*, Vol. 28 No. 1, pp. 1-26, doi: 10.1080/10920277.2022.2162924.

Tibshirani, R. (1996), "Regression shrinkage and selection via the lasso", *Journal of the Royal Statistical Society Series B: Statistical Methodology*, Vol. 58 No. 1, pp. 267-288, doi: 10.1111/j.2517-6161.1996.tb02080.x.

Trawiński, B., Smętek, M., Telec, Z. and Lasota, T. (2012), "Nonparametric statistical analysis for multiple comparison of machine learning regression algorithms", *International Journal of Applied Mathematics and Computer Science*, Vol. 22 No. 4, pp. 867-881, doi: 10.2478/v10006-012-0064-z.

Turvey, C.G. (2001), "Weather derivatives for specific event risks in agriculture", *Applied Economic Perspectives and Policy*, Vol. 23 No. 2, pp. 333-351, doi: 10.1111/1467-9353.00065.

Turvey, C.G., Du, J., He, Y. and Ortiz-Bobea, A. (2021), "A vulnerability index for priority targeting of agricultural crops under a changing climate", *Climatic Change*, Vol. 166 No. 34, 34, doi: 10.1007/s10584-021-03135-8.

Vabalas, A., Gowen, E., Poliakoff, E. and Casson, A.J. (2019), "Machine learning algorithm validation with a limited sample size", *PloS One*, Vol. 14 No. 11, e0224365, doi: 10.1371/journal.pone.0224365.

Vallentin, C., Harfenmeister, K., Itzerott, S., Kleinschmit, B., Conrad, C. and Spengler, D. (2022), "Suitability of satellite remote sensing data for yield estimation in northeast Germany", *Precision Agriculture*, Vol. 23 No. 1, pp. 52-82, doi: 10.1007/s11119-021-09827-6.

Vedenov, D.V. and Barnett, B.J. (2004), "Efficiency of weather derivatives as primary crop insurance instruments", *Journal of Agricultural and Resource Economics*, Vol. 29 No. 3, pp. 387-403.

Verstegen, H., Köneke, O., Korzun, V. and von Broock, R. (2014), "The world importance of barley and challenges to further improvements", in Kumlehn, J. and Stein, N. (Eds), *Biotechnological Approaches to Barley Improvement*, Springer, Berlin, Heidelberg, pp. 3-19.

Vroege, W., Bucheli, J., Dalhaus, T., Hirschi, M. and Finger, R. (2021), "Insuring crops from space: the potential of satellite-retrieved soil moisture to reduce farmers' drought risk exposure", *European Review of Agricultural Economics*, Vol. 48 No. 2, pp. 266-314, doi: 10.1093/erae/jbab010.

Wahid, A., Gelani, S., Ashraf, M. and Foolad, M.R. (2007), "Heat tolerance in plants: an overview", *Environmental and Experimental Botany*, Vol. 61 No. 3, pp. 199-223, doi: 10.1016/j.envexpbot.2007.05.011.

Werner, W.J., Sanderman, J. and Melillo, J.M. (2020), "Decreased soil organic matter in a long-term soil warming experiment lowers soil water holding capacity and affects soil thermal and hydrological buffering", *Journal of Geophysical Research: Biogeosciences*, Vol. 125 No. 4, e2019JG005158, doi: 10.1029/2019jg005158.

Woodard, J.D. and Garcia, P. (2008), "Weather derivatives, spatial aggregation, and systemic risk: implications for reinsurance hedging", *Journal of Agricultural and Resource Economics*, Vol. 33 No. 1, pp. 34-51.

Zhang, J., Tan, K.S. and Weng, C. (2019), "Index insurance design", *ASTIN Bulletin: The Journal of the IAA*, Vol. 49 No. 2, pp. 491-523, doi: 10.1017/asb.2019.5.

Zou, J., Odening, M. and Okhrin, O. (2023), "Plant growth stages and weather index insurance design", *Annals of Actuarial Science*, Vol. 17 No. 3, pp. 438-458, doi: 10.1017/s1748499523000167.

Appendix

The supplementary material for this article can be found online.

Corresponding author

Jing Zou can be contacted at: jing.zou@mailbox.tu-dresden.de