ESTIMATION OF THE FARM-LEVEL YIELD-WEATHER-RELATION USING MACHINE LEARNING

Lorenz Schmidt, Martin Odening, Johann Schlanstein, Matthias Ritter

matthias.ritter@agrar.hu-berlin.de

Humboldt-Universität zu Berlin, Lebenswissenschaftliche Fakultät, Thaer-Institut für Agrar- & Gartenbauwissenschaften, Department für Agrarökonomie, Unter den Linden 6, D-10099 Berlin

Acknowledgment

The authors are grateful to *Vereinigte Hagelversicherung* and *LandData GmbH* for providing the dataset. Financial support from the Federal Ministry of Education and Research (BMBF) through the project KlimALEZ is gratefully acknowledged.

2021

Vortrag anlässlich der 61. Jahrestagung der GEWISOLA

(Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V.)

Transformationsprozesse im Agrar- und Ernährungssystem: Herausforderungen für die Wirtschafts- und Sozialwissenschaften,

22. bis 24. September 2021

ESTIMATION OF THE FARM-LEVEL YIELD-WEATHER-RELATION USING MACHINE LEARNING

Abstract

Weather is a pivotal factor for crop production as it is highly volatile and can hardly be controlled by farm management practices. Since there is a tendency towards increased weather extremes in the future, understanding the weather-related yield factors becomes increasingly important not only for yield prediction, but also for the design of insurance products that mitigate financial losses for farmers. In this study, an artificial neural network is set up and calibrated to a rich set of farm-level wheat yield data in Germany covering the period from 2003 to 2018. A nonlinear regression model, which uses rainfall, temperature, and soil moisture as explanatory variables for yield deviations, serves as a benchmark. The empirical application reveals that the gain in estimation precision by using machine learning techniques compared with traditional estimation approaches is quite substantial and that the use of regionalized models and high-resolution weather data improve the performance of ANN.

Keywords

Yield Prediction, Machine Learning, Weather Risk, Risk Management, Index Insurance.

1 Introduction

Understanding yield variability is essential for agricultural risk management on a sectoral as well as on a farm level. Crop yields depend on a variety of factors including soil and weather conditions, fertilizer, and pest control. Among these factors, weather is pivotal because – in contrast to other production factors – it is highly volatile and can hardly be controlled by farm management practices. In fact, extreme weather events lead to harvest failures and thus threaten food security all over the world (WHEELER and BRAUN, 2013). Since there is a tendency towards increased weather extremes in the future, understanding the weather-related yield factors will become increasingly important not only for yield prediction, but also for the design of insurance products that mitigate financial losses for farmers. Actually, weather-based insurance products, such as index insurance and weather derivatives, have been propagated as a promising alternative to classical crop insurance (BARNETT and MAHUL, 2007). A main finding of the vast literature is that the effectiveness of weather-based insurance hinges on a high correlation between actual yields and the insured weather event (WOODARD and GARCIA, 2008). However, the relationship between weather and crop yield is complex, which challenges the design of appropriate weather indices. Firstly, several weather variables have to be considered simultaneously, particularly precipitation, temperature, and wind. Secondly, these variables interact in a highly nonlinear way (SCHLENKER and ROBERTS, 2009). Finally, not only their levels but also their temporal distributions affect crop yields (MUSSHOFF et al., 2011).

Two approaches have been mainly used for modeling the weather-yield nexus: Firstly, crop growth models that rest on biological and physical relations and simulate the dynamics of water, nitrogen, carbon, and other yield determinants in a specific soil context taking into account phenological stages and plant requirements (e.g. ASSENG, 2004). Alternatively, statistical methods, regressions models in particular, have been employed to estimate crop yields as a function of weather variables (see Section 2 for a detailed literature review). These methods are mainly data driven and do not strive for an identification of causal relations. In this paper, we focus on statistical approaches, as they are most common in the context of weather insurance. MUSSHOFF et al. (2011) show that a trade-off exists between the regression model's simplicity and basis risk, i.e., the yield variation that cannot be explained by weather variables. Several directions have been suggested to improve the fit of statistical yield models, including nonlinear regression or quantile regression (CONRADT et al., 2015). More recently, machine learning techniques have been applied to yield modelling (e.g. KHAKI and WANG, 2019). The strength of this approach compared with traditional statistical methods arises from its flexibility in capturing complex functional relations and the capability to handle large data sets. This is particularly useful because it allows to consider weather variables with high temporal resolution, e.g., daily precipitation or temperature.

Against this backdrop, the objective of our paper is to explore the potential of machine learning for estimating the relationship between crop yield and weather conditions on a farm level. More specifically, we want to investigate two hypotheses: On the one hand, we conjecture that machine learning allows a better fit to yield data compared with traditional regression models due to its flexibility. On the other hand, we hypothesize that disaggregated weather data contain more information compared with aggregated weather variables, which allow improving the estimation of crop yields. We test these hypotheses for a large set of farm-level wheat yields. Our data set contains 68,944 yield observations in total covering many production regions in Germany over an observation period of 16 years. The use of individual farm yields avoids the underestimation of yield volatility that arises from usage of aggregated data, such as county yields (POPP et al., 2005).

The remainder of this paper is structured as follows. Section 2 provides a literature review of standard statistical as well as machine learning approaches to estimate the weather-yield relation. In Section 3, we present details on the neural network applied in this study and introduce a baseline model that is used as a benchmark. Section 4 contains the empirical application to German farm-level data. Section 5 concludes.

2 Literature Review

Before the use of machine learning, the weather-yield relation was analyzed using traditional statistical approaches. TEIGEN and THOMAS (1995) study the relation for state-level yield for the period 1950–1994 and can explain 90 % of the yield variation in most cases. This high percentage, however, can mostly be traced back to the time trend and not to the weather variables themselves (VEDENOV and BARNETT, 2004). For the application of weather derivatives to agriculture, TURVEY (2001) estimates the linear dependency of county yields of corn, soybean, and hay on cumulative rainfall and cumulated degree days in Oxford County, Ontario, for the period 1935–1996. The best fit amounts to an R^2 of 0.33. Also in the context of weather derivatives, VEDENOV and BARNETT (2004) apply more complex non-linear models to estimate the relation between U.S. district-level yields in 1972–2001 and temperature and precipitation. With data-driven combinations of the weather variables and derived indices, they achieve an R^2 between 35 % and 87 %. VROEGE et al. (2021) assess the potential of drought risk management with soil moisture data from satellites and weather stations for 89 farms in Eastern Germany. They apply a quantile regression and find that the risk exposure of farmers could be reduced significantly with new insurance products based on soil moisture. Besides weather risk management, another purpose of the statistical modelling of the yield-weather relationship is the prediction of climate change impacts. Seminal papers in this context are SCHLENKER and ROBERTS (2006, 2009), who combine a county-level data set for U.S. maize yield with daily temperature observations and observe non-linear weather effects on yields, and SCHLENKER and LOBELL (2010), who apply different specifications of the weather variables (linear, quadratic, and piece-wise linear) and find robust negative effects of climate change on agriculture in Africa. On a country-level, LOBELL et al. (2011) regress yield outcomes on linear and squared monthly temperature and precipitation. It turns out, however, that the largest share of the explained variation comes from the country-specific intercepts and the quadratic time trend and not the weather variables. To detect spatio-temporal patterns in the yield-weather relation, TRNKA et al. (2016) use data for ten countries and two regions in Europe in the period 1901–2012 for wheat and barley. Additional to the classical weather variables, they apply drought indicators, frost days, potential evapotranspiration, and water vapor pressure deficit, but achieve a rather poor fit with an adjusted R^2 for wheat between 0.00 and 0.71 and an RMSE between 65 % and 130 % also when looking at subperiods. Nevertheless, they find an increasing influence of climatic variables in the later years.

All these studies show how difficult it is to explain the yield-weather relation using classical statistical approaches. Hence, a lot of hope is put in the use of machine learning and the increased computational power, which allow a more sophisticated analysis of the relation. VAN KLOMPENBURG et al. (2020) conduct a systematic literature review and identify 50 studies since 2008 that use machine learning for crop yield modelling. Explanatory variables are mostly related to weather, but also other features such as field management or nutrients are considered. For example, MATSUMARA et al. (2015) predict the maize yield in Chilin province, China, based on weather variables and fertilizer usage using a multi-layer perceptron with one hidden layer and compare the results with those of a linear regression model. Whereas the artificial neural network clearly outperforms the linear regression model, the predictive performance can mainly be traced back to the fertilizer and not to the weather variables. JEONG et al. (2016) apply random forests to global wheat yield raster data in 2000, U.S. county-level maize grain yield 1984–2013, and potato tuber and maize silage yield data for over 1,000 points in the Northeastern U.S. in selected years. They achieve an RMSE between 6 % and 14 %, which clearly outperforms a multiple regression model (RMSE between 14 % and 49 %). Also with random forests, EVERINGHAM et al. (2016) aim to predict regional sugarcane yields at Tully, Australia, at different time points up to a year before harvest to optimize fertilizer usage. The shorter the forecast horizon, the more important become variables such as rainfall and temperature range, and up to 79 % of the variability can be explained. Using a semiparametric version of a deep neural network, CRANE-DROESCH (2018) models county-level yield in the U.S. Midwest from 1979 to 2016 in dependence of daily weather variables such as precipitation, temperature, humidity, wind speed, and radiation. It turns out that while the semiparametric model is best (with the largest effect of a time variable), the fully nonparametric neural network performs much worse than an OLS regression. In a crop modelling challenge, KHAKI and WANG (2019) as one of the winning teams achieve an RMSE of 12 % with a deep neural network when predicting the yield performance of maize hybrids at over 2,000 locations in the U.S. They find considerable effects of solar radiation, temperature, and precipitation. Some studies also use remote sensing data and derived indices such as the Normalized Difference Vegetation Index (NDVI) or Enhanced Vegetation Index (EVI) (e.g. PANTAZI et al., 2016; JOHNSON et al., 2016; FERNANDES et al., 2017; SUN et al., 2019; WOLANIN et al., 2020).

Applications of machine learning methods for the estimation of the yield-weather relation in Germany, however, are rare. PAUDEL et al. (2021) design a workflow for large-scale crop yield forecasting at different steps between planting and harvesting and apply it to the Netherlands, Germany, and France. For whole Germany, they achieve a normalized RMSE between 7 % and 17 % at the end of season, which is much larger than the corresponding values from predictions by the European Commission's MARS Crop Yield Forecasting System (MCYFS). On a county-level, WEBBER et al. (2020) combine support vector machines and process-based modelling using data on weather, soil, and crop phenology to explain yield failures. Their model, however, was not able to capture the losses in 2018, an exceptionally dry year in Central Europe (TORETI et al., 2019).

It can be concluded that the use of machine learning does not automatically lead to better results and requires a careful specification. Moreover, crop yield modelling is usually done at a larger scale, so that our study is – to our best knowledge – the first in modelling crop yield based on a rich farm-level data set in Germany.

3 Methodology

3.1 Baseline Model

Our baseline model, which serves as a benchmark for the neural network model, is a multiple regression model. As the dependent variable, we use the deviation of the yield from the farm yield average in the training data measured in dt/ha. By subtracting the farm-specific mean, we remove constant location- or farmer-specific factors influencing the yield to reduce the omitted variable bias of the model. Following VEDENOV and BARNETT (2004) and VROEGE et al. (2021), we use the average temperature, the total precipitation, and the average soil moisture as independent variables. All weather variables are calculated as monthly values for April, May, and June, which represent the growing period. As in VEDENOV and BARNETT (2004), we additionally apply squares and same-month interactions of these variables to allow for a non-linear relation. The regression model can be defined as follows:

$$\Delta y_i = \beta_0 + \sum_{k=\text{April, May, June}} \beta_{1k} T_{ki} + \beta_{2k} P_{ki} + \beta_{3k} M_{ki} + \beta_{4k} T_{ki}^2 + \beta_{5k} P_{ki}^2 + \beta_{6k} M_{ki}^2 + \beta_{7k} T_{ki} P_{ki} + \beta_{8k} T_{ki} M_{ki} + \beta_{9k} P_{ki} M_{ki} + \beta_{10k} T_{ki} P_{ki} M_{ki} + \epsilon_i \quad (1)$$

where Δy_i denotes the yield deviation for farm i and T_{ki}, P_{ki}, and M_{ki} the values of the weather variables temperature, precipitation, and soil moisture, respectively, at farm i in month k (April, May, June). The βs denote the coefficients to be estimated and ϵ_i the error term. To estimate the model parameters, we use the ordinary least square (OLS) method.

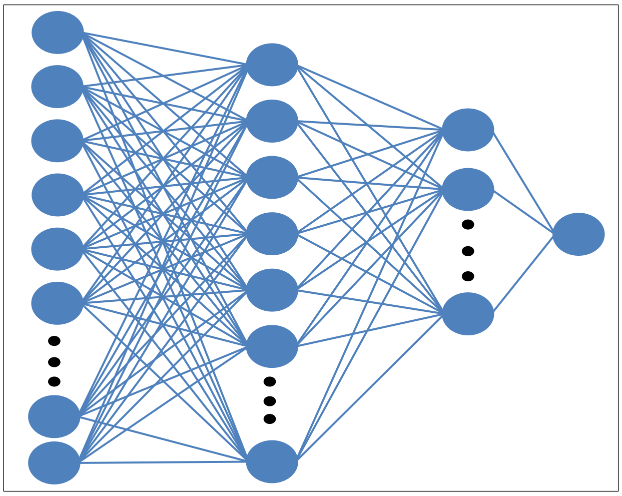
3.2 Neural Network

Second, we apply a neural network (NN) to estimate the weather-yield relationship based on the same dependent variable as in the previously described model. NN with at least two hidden layers are able to recreate any form of a mathematical model which is in line with the non-linear relationship between weather and crop yields (SHARMA et al., 2020). This model is illustrated in Figure 1. While setting up and training a NN, hyperparameter tuning is essential, but there is no strict method for selecting and tuning hyperparameters so far. In our study, we develop an NN of two hidden layers while performing grid search on a search space (Table A1) with Tune as platform (LIAW et al., 2018). We use the setting of hyperparameters with the lowest RMSE on the validation set. The search space for the grid search included learning rate, batch size, and number of neurons per hidden layer as hyperparameters. Since we are facing a regression problem, we have one neuron in the output layer. For training the model, we use stochastic gradient descent and the Adam optimizer (KINGMA and BA, 2014). To account for the non-linear relationship between weather and crop yields, we decided in line with SHARMA et al. (2020) for a non-linear activation function and to use the ReLU function (rectifier linear unit). It can be written as $g(x) = \max(0, x)$. The activation function is used for all neurons in the layer except for the output layer. The NN was implemented in Python using the PyTorch library and trained on a Linux engine (PASZKE et al., 2019). Before training the NN, the features were normalized.

Every data set is split into three subsets: the training data set, validation data set as well as the test data set (NORDHAUSEN, 2009). The training data set is used to adjust the weights and to train the model. The validation set is used to evaluate the results of the grid search and to choose for the hyperparameter settings with the lowest RMSE. In the end, the independent test set is used for evaluating the out-of-sample performance of the model.

To analyse the effects of aggregated and disaggregated weather data, we feed the neural network with daily or monthly weather data. Therefore, we apply the monthly mean of soil moisture and temperature and the monthly sum of precipitation.

Figure 1: Neural Network with two fully connected hidden layers



4 Empirical Application

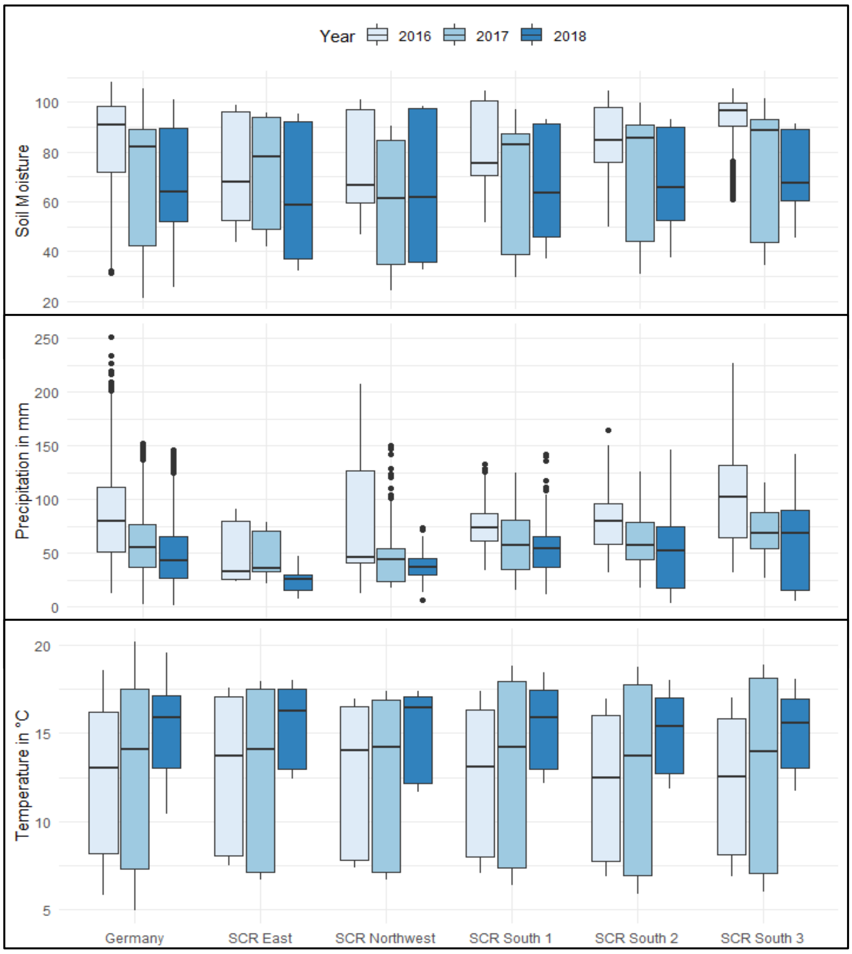
4.1 Study Region and Data

In the empirical application, we use a data set with the annual winter wheat yield from 4,309 German farms in 2003–2018, measured in deciton/hectare (dt/ha). In total, the data set consists of 68,944 observations. Germany is advantageous for investigating the effects of drought on winter wheat since only about 2.7 % of the agricultural area in Germany is irrigated (SCHIMMELPFENNIG et al., 2018). The data set was provided by a financial accounting firm and an insurance company that has asked farmers for voluntary information on harvest quantity and area. The farms are spread across Germany with a majority in the south. The exact locations have been anonymized, but the municipalities in which the farms are located are provided.

Since the yield data were collected through a voluntary survey, some inaccuracies may exist. For example, some respondents have entered the same value repeatedly although the years differ significantly in their weather conditions. Here, we suspect that respondents have not entered realistic values and therefor remove farms where at least half of the yields were identical. To correct outliers from inaccuracies in the data collecting process, we identify the 1% percentile and the 99% percentile with the corresponding farms and deleted those farms from the data set. This is done for each study region individually.

The weather data are provided by the Climate Data Center of *Deutscher Wetterdienst* (DWD). We utilize daily precipitation, daily temperature, and daily soil moisture data spanning from 2003 to 2018 corresponding to the agricultural data. The temperature is an average of 24-hourly values and is measured in °C two meters above the surface. The amount of precipitation is measured in mm. Soil moisture data are estimated by the water balanced model AMBAV (agrometeorological model to calculate the current evaporation) (LÖPMEIER, 1994). Since we do not know the exact locations of the farms, we connect the yield data with the weather data via the respective municipality. The DWD interpolates temperature and precipitation data coming from around 300 weather stations to a 1 km x 1 km grid based on the interpolation method by FREI (2014). With 2018, we have an exceptional dry year in central Europe in our test set (TORETI et al., 2019). It can also be observed in our weather data for all German locations (Figure 2 left) that 2018 had more extreme conditions compared to 2016 and 2017. While there was a generally higher level of the temperature, 2018 had both normal and particularly low soil moisture levels. In the case of precipitation, a generally lower level can be observed in 2018.

Figure 2: Monthly weather values for all farm locations (Germany) as well as for selected soil-climate-regions (SCRs)



Since the conditions to farm agricultural land vary greatly in the different regions of Germany, we also split our data set into regions with comparable soil and weather conditions. The chambers of agriculture of the federal states and the federal biological research centre for agriculture and forestry have agreed on the soil-climate-region (SCR) classification. A clustering procedure was used to combine municipalities with similar characteristics in terms of soil quality, temperature, and precipitation to larger areas, which have relatively homogeneous conditions for agricultural production (ROßBERG et al., 2007). In our analysis, we first use the entire data set (Germany), then the three SCRs with the largest number of farms in our data set (SCR South 1, SCR South 2, SCR South 3) as well as one SCR in north-western Germany (SCR Northwest) and one in eastern Germany (SCR East).[1] In Figure 3, the location of the selected soil-climate-regions can be seen. The descriptive statistics for the cleaned yield data sets of these subgroups are depicted in Table 1, the monthly weather variables for the SCRs in Figure 2.

[1] The exact names of the soil-climate-regions are: SCR South 1: BKR113 – 'Nordwestbayern-Franken'; SCR South 2: BKR114 – 'Albflächen und Ostbayerisches Hügelland'; SCR South 3: BKR115 – 'Tertiär-Hügelland Donau-Süd'; SCR Northwest: BKR147 – 'mittleres Niedersachsen, nordöstliches NRW'; SCR East: BKR108 – 'Lößböden in den Übergangslagen (Ost)'

Figure 3: Soil-climate-regions (SCR) in Germany (ROßBERG et al., 2007) and presentation of the SCRs considered in this study

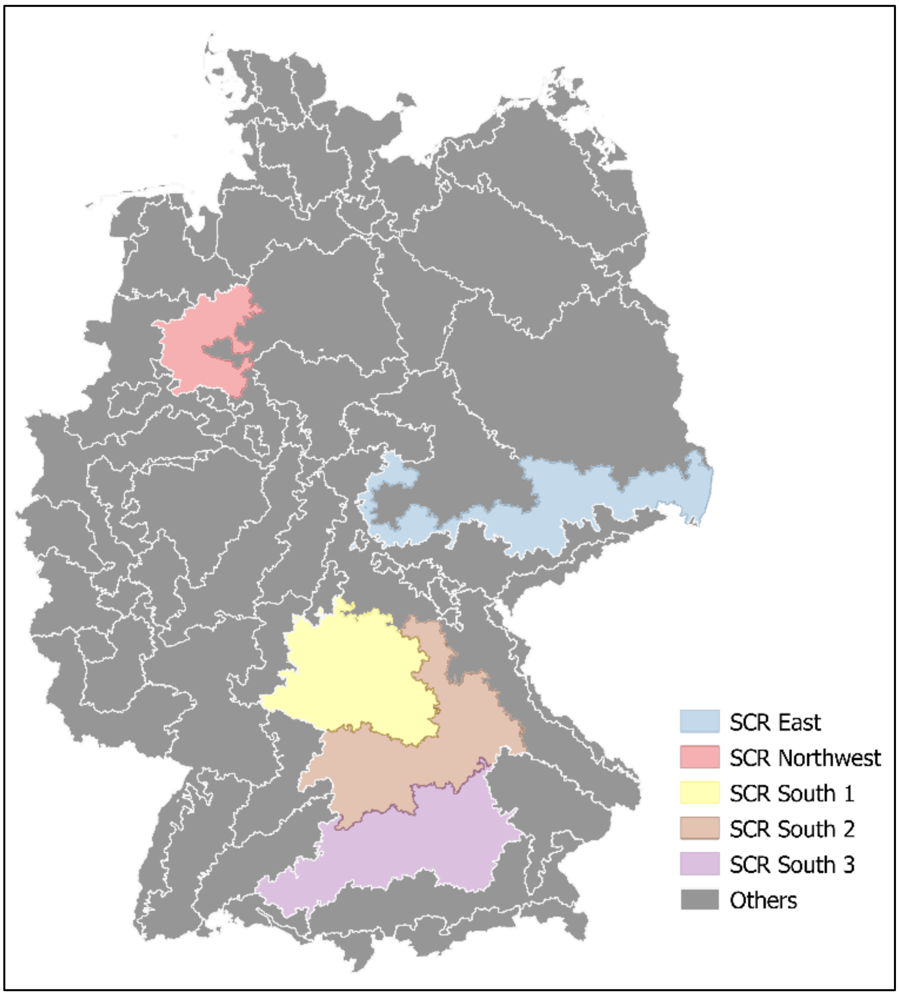


Table 1: Descriptive statistics of the yield data (dt/ha) for whole Germany and the considered soil-climate-regions (SCR)

	# farms	# obs.	Mean	St. Dev.	Min.	25 %	50 %	75 %	Max.
Germany	3,344	53,504	74.20	13.57	26.59	65.86	75.09	83.02	113.58
SCR South 1	373	5,968	67.57	14.27	21.54	59.48	69.65	78.21	101.69
SCR South 2	482	7,712	73.96	12.17	32.21	66.84	74.99	81.21	114.25
SCR South 3	394	6,304	77.64	12.06	31,26	71.11	78.84	85.00	113.44
SCR Northwest	97	1,552	80.11	12.83	37.65	72.63	80.28	89.01	136.16
SCR East	7	112	79.33	13.50	51.48	70.19	79.91	89.84	106.81

4.2 Results

We start with the results for the whole data set before we take a closer look into specific soil-climate-regions to examine regional and temporal differences. The applied data set is split by years into training data (2003–2012), validation data (2013–2015), and testing data (2016–2018) to create a realistic scenario as we would have it in an insurance application. Whereas the training and validation data are used to specify and estimate the model, the performance of the different models is eventually compared based on the independent test set. Even if this split is not necessary for the baseline model, we also apply it to ensure comparability across the models. For each machine learning model, a separate grid search was performed to improve the performance. While in some regions, only marginal improvements could be achieved, the RMSE could be reduced by about 40 % in other regions through grid search. Although this method is very resource- and time-intensive, it offers the possibility to achieve comparable results for different trials. The best performing hyperparameter configurations are shown in Table A2. During the training of the model, it turned out that overfitting already occurred after a few iterations depending on the region. With the feature normalization, it was tried to counteract this and to enhance the performance of the model. This also accounts for the different dimension in the input variables.

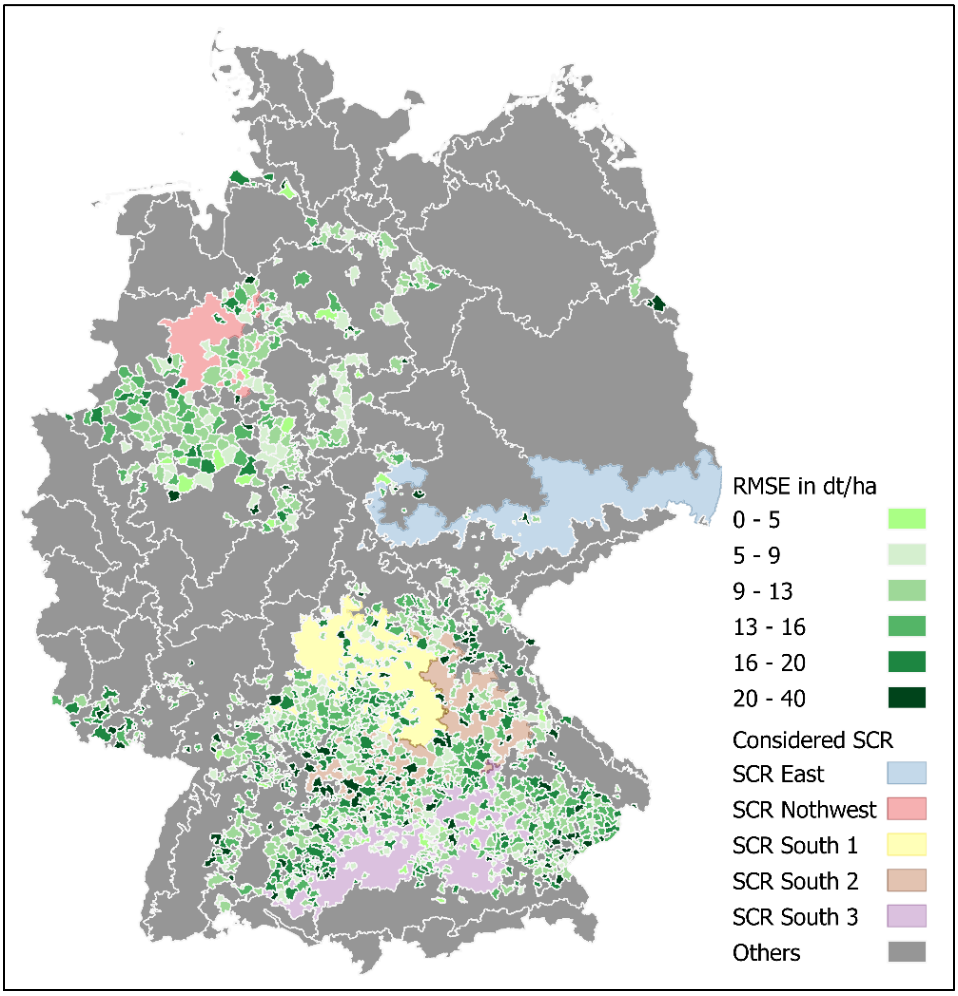
Table 2 depicts the RMSE for models using all farms of the data set (Germany). The RMSE for the testing data with the baseline model amounts to 13.06 dt/ha. Compared to an average yield of 74.20 dt/ha of the whole study period, this error is substantial. The RMSE for the training data of 10.23 dt/ha, however, shows that the regression model also cannot explain a much larger share of the yield deviations in-sample. Applying the neural network with monthly data, surprisingly even further increases the RMSE of the test set to 14.44 dt/ha. The switch to daily weather variables reduces the RMSE for the testing data to 12.38 dt/ha. Evaluating the performance of the models for the five SCRs reveals that the NN with monthly data performs worst in all southern SCRs whereas the NN with monthly data outperforms the baseline model in SCR East and SCR Northwest (Table 2). The NN with daily data constantly performs best, even though sometimes only with small differences.

In Figure 4, the location of the municipalities in which the farms are located can be observed. As mentioned before, most of the farms are located in the south of Germany and some of them are in the west and east. The map also depicts the RMSE for each municipality based on the same model as above showing a large range.

Table 2: Root mean squared error (RMSE) for baseline and NN models based on all farms, evaluated for whole Germany and the five selected SCRs

Iterations = 100	Data set	Baseline Model	NN Monthly Data	NN Daily Data
Germany	Training	10.2322	7.9910	8.3717
	Validation	12.2169	12.7779	12.1812
	Testing	13.0645	14.4394	12.3756
SCR East	Testing	16.0009	14.9553	14.7233
SCR Northwest	Testing	13.9559	12.7541	12.3952
SCR South 1	Testing	13.3573	14.2376	12.4437
SCR South 2	Testing	13.1562	15.5831	12.5938
SCR South 3	Testing	12.4398	14.9825	12.4165

Figure 4: RMSE per municipality for NN based on monthly data



We want to investigate two potential explanations for the rather high error in general: First, using the same model for all farms in Germany might be inadequate given the large heterogeneity of farming and weather conditions. Hence, we will split the data into subsets using the aforementioned SCRs and estimate a separate model for each SCR. Second, the particularly dry year 2018 might be responsible for the large error, so that we will have a closer look at the performance of the models in the single years of the testing set.

We start with splitting the data sets into the selected SCRs and estimating SCR-specific models. The results in Table 3 strongly differ between the three southern SCRs and the other two SCRs. Regarding the baseline model, the southern SCRs have an RMSE for the testing data between 12.25 and 13.07 dt/ha, which is not too different from the baseline model results from Table 2 for one model for all farms (between 12.43 and 13.36 dt/ha). The NN with monthly data does not change the performance substantially whereas the NN with daily data reduces the RMSE to between 10.72 and 11.97 dt/ha. The latter outperforms the model based on all farms with an RMSE between 12.42 and 12.59 dt/ha, so that estimating separate SCR-specific models seems beneficial.

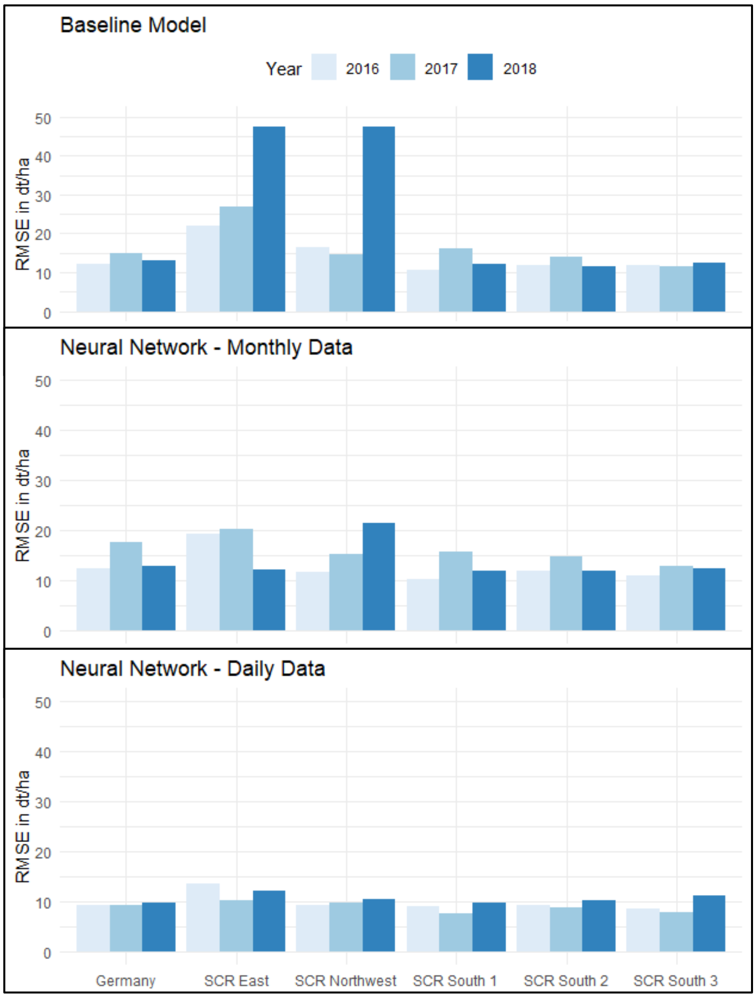
On the other side, the results for SCR East and SCR Northwest show a different picture. The RMSE for the baseline model increases to 28.38 (SCR Northwest) and 38.90 dt/ha (SCR East) and for the NN with monthly data to 16.68 and 17.67 dt/ha, respectively. These errors are much larger compared to those based on one model for all farms (between +3.92 and +22.90). Only the NN with daily data shows comparable results, with a clear decrease in the RMSE for SCR East (–2.60) and a slight increase for SCR Northwest (+0.86). It turns out that estimating SCR-specific models can substantially worsen the results whereas only the NN with daily data seems to have a robust performance. By using daily weather data, the NN has far more parameters that can be trained compared to the NN with monthly data and so it is able to better capture certain weather events. A substantial difference between the southern SCRs and the other two is the number of farms and hence the number of observations in the data set. The southern SCRs include between 373 and 482 farms whereas the other two consist only of 97 (SCR Northwest) or even 7 farms (SCR East). Given the size of the data sets, the results may lead to the conclusion that the NN can be improved by using regionalized data sets containing farms with similar soil and climatic conditions, but that they must be of a certain size to benefit from these similarities.

To examine the temporal dimension of the RMSE and the influence of the particularly dry year 2018, we compare the RMSE for each year separately for one model for all farms (Germany) and the five SCR-specific models (Figure 5). The RMSE for the baseline model is particularly high in SCR East and SCR Northwest in 2018. From the monthly weather values in Figure 2, however, it cannot be concluded that 2018 was a special year only in these regions, so that the exact reason for the high RMSE remains unclear. The performance of the NN based on monthly data also differs between the three years although with a smaller range. The NN with daily data does not only lead to the smallest RMSE, but its performance also varies little between the three years. This shows that the NN based on daily data yields to stable results, even in particularly dry years such as 2018.

Table 3: RMSE for five SCR-specific baseline and NN models

Iterations = 100	Data set	Baseline Model	NN Monthly Data	NN Daily Data
SCR East	Training	12.9636	7.5726	8.0731
	Validation	37.8269	10.7268	13.2185
	Testing	38.8951	17.6671	12.1200
SCR Northwest	Training	10.3422	8.1497	8.9967
	Validation	17.4650	14.8219	11.0313
	Testing	28.3832	16.6757	13.2549
SCR South 1	Training	12.7603	7.8476	7.5579
	Validation	14.3575	9.6640	10.0686
	Testing	13.0683	12.8757	10.7178
SCR South 2	Training	10.4563	7.7469	7.6343
	Validation	11.9801	11.4379	10.7674
	Testing	12.4357	12.9929	11.9735
SCR South 3	Training	9.8270	7.8283	7.6072
	Validation	12.9103	12.5030	10.2084
	Testing	12.2540	12.2230	11.2918

Figure 5: RMSE by year of the testing set for one model for all farms (Germany) and SCR-specific models



5 Conclusions

In this paper, we explore the potential of machine learning techniques for improving the estimation of weather-induced yield losses. We specify an artificial neural network and calibrate it to a rich set of farm-level wheat yield data in Germany covering the period from 2003 to 2018. A nonlinear regression model, which uses rainfall, temperature, and soil moisture as explanatory variables for yield deviations, serves as a benchmark. Our empirical application reveals that the gain in estimation precision by using machine learning techniques compared with traditional estimation approaches is quite substantial. The reduction of the RMSE on the test data amounts to 30 percent on average for the regionalized models. While the use of daily weather data instead of monthly weather data lead to a significant improvement of the model fit for all models, the use of regionalized models is only beneficial if the region is of a certain size. It is noteworthy that even for the best fitting ANN, the level of the RMSE amounts to more than 10 dt/ha and is quite high relative to the average wheat yield level. This reveals that a considerable part of the yield variability on a farm level is unsystematic and hard to predict by statistical methods or the use of "big weather data".

This finding has important implications for the design of weather-index based insurance because it documents that a rather high level of basis risk remains with insured farms if insurance products are based on a general weather-yield relationship. This suggests the use of other indices, such as area yields, as an underlying for index-based insurance. Our results, however, should be considered with caution because they are only a first attempt to tap the full potential of machine learning in this context. Future research should use models with flexible model structures, e.g., convolutional neural networks or locally connected layers, to better estimate the meteorological factors affecting yield. We also propose the application of neural networks with high-resolution data to other crops and regions to generalize the findings of our study.

6 Appendix

Table A1: Search Space for hyperparameter grid search

Learning Rate	0.001	0.002	0.004	0.0065	0.008	0.016	0.032	0.064	0.08	0.12
Batch Size	8	16	24	32	40	48				
# Neurons /Hidden-Layer	40	45	50	55	60	65	70	75		

Table A2: Hyperparameter configurations after grid search

		Batch Size	Learning Rate	# Neurons /Hidden-Layer
Germany	Monthly	8	0.002	65
	Daily	8	0.001	65
SCR East	Monthly	48	0.001	65
	Daily	48	0.12	50
SCR Northwest	Monthly	8	0.001	65
	Daily	8	0.016	75
SCR South 1	Monthly	8	0.001	60
	Daily	8	0.008	70
SCR South 2	Monthly	8	0.001	40
	Daily	24	0.032	75
SCR South 3	Monthly	8	0.001	65
	Daily	8	0.032	70

7 References

ASSENG, S. (2004): Wheat crop systems. A simulation analysis. CSIRO Publishing, Melbourne.

BARNETT, B. J. and O. MAHUL (2007): Weather Index Insurance for Agriculture and Rural Areas in Lower-Income Countries. In: American Journal of Agricultural Economics 89 (5): 1241–1247.

CONRADT, S., R. FINGER and R. BOKUSHEVA (2015): Tailored to the extremes: Quantile regression for index-based insurance contract design. In: Agricultural Economics 46 (4): 537–547.

CRANE-DROESCH, A. (2018): Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. In: Environmental Research Letters 13 (11): 114003.

EVERINGHAM, Y., J. SEXTON, D. SKOCAJ and G. INMAN-BAMBER (2016): Accurate prediction of sugarcane yield using a random forest algorithm. In: Agronomy for Sustainable Development 36 (2).

FERNANDES, J. L., N. F. F. EBECKEN and J. C. D. M. ESQUERDO (2017): Sugarcane yield prediction in Brazil using NDVI time series and neural networks ensemble. In: International Journal of Remote Sensing 38 (16): 4631–4644.

FREI, C. (2014): Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances. In: International Journal of Climatology 34 (5): 1585–1605.

JEONG, J. H., J. P. RESOP, N. D. MUELLER, D. H. FLEISHER, K. YUN, E. E. BUTLER, D. J. TIMLIN, K.-M. SHIM, J. S. GERBER, V. R. REDDY and S.-H. KIM (2016): Random Forests for Global and Regional Crop Yield Predictions. In: PloS one 11 (6): e0156571.

JOHNSON, M. D., W. W. HSIEH, A. J. CANNON, A. DAVIDSON and F. BÉDARD (2016): Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods. In: Agricultural and Forest Meteorology 218-219: 74–84.

KHAKI, S. and L. WANG (2019): Crop Yield Prediction Using Deep Neural Networks. In: Frontiers in plant science 10: 621.

KINGMA, D. P. and J. BA (2014): Adam: A Method for Stochastic Optimization. In: https://arxiv.org/pdf/1412.6980.

LIAW, R., E. LIANG, R. NISHIHARA, P. MORITZ, J. E. GONZALEZ and I. STOICA (2018): Tune: A Research Platform for Distributed Model Selection and Training. In: https://arxiv.org/pdf/1807.05118.

LOBELL, D. B., W. SCHLENKER and J. COSTA-ROBERTS (2011): Climate trends and global crop production since 1980. In: Science (New York, N.Y.) 333 (6042): 616–620.

LÖPMEIER, F.-J. (1994): Berechnung der Bodenfeuchte und Verdunstung mittels agrarmeteorologischer Modelle. In: Zeitschrift für Bewässerungswirtschaft (29): 157–167.

MATSUMARA, K., C. F. GAITAN, K. SUGIMOTO, A. J. CANNON and W. W. HSIEH (2015): Maize yield forecasting by linear regression and artificial neural networks in Jilin, China. In: The Journal of Agricultural Science 153 (3): 399–410.

MUSSHOFF, O., M. ODENING and W. XU (2011): Management of climate risks in agriculture–will weather derivatives permeate? In: Applied Economics 43 (9): 1067–1077.

NORDHAUSEN, K. (2009): The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition by Trevor Hastie, Robert Tibshirani, Jerome Friedman. In: International Statistical Review 77 (3): 482.

PANTAZI, X. E., D. MOSHOU, T. ALEXANDRIDIS, R. L. WHETTON and A. M. MOUAZEN (2016): Wheat yield prediction using machine learning and advanced sensing techniques. In: Computers and Electronics in Agriculture 121: 57–65.

PASZKE, A., S. GROSS, F. MASSA, A. LERER, J. BRADBURY, G. CHANAN, T. KILLEEN, Z. LIN, N. GIMELSHEIN, L. ANTIGA, A. DESMAISON, A. KOPF, E. YANG, Z. DEVITO, M. RAISON, A. TEJANI, S. CHILAMKURTHY, B. STEINER, L. FANG, J. BAI and S. CHINTALA (2019): PyTorch: An Imperative Style, High-Performance Deep Learning Library. In: H. Wallach et al. (Eds.): Advances in Neural Information Processing Systems 32. Curran Associates, Inc: 8024–8035.

PAUDEL, D., H. BOOGAARD, A. de WIT, S. JANSSEN, S. OSINGA, C. PYLIANIDIS and I. N. ATHANASIADIS (2021): Machine learning for large-scale crop yield forecasting. In: Agricultural Systems 187: 103016.

POPP, M., M. RUDSTROM and P. MANNING (2005): Spatial Yield Risk Across Region, Crop and Aggregation Method. In: Canadian Journal of Agricultural Economics (53): 103–115.

ROßBERG, D., V. MICHEL, R. GRAF and R. NEUKAMPF (2007): Definition von Boden-Klima-Räumen für die Bundesrepublik Deutschland. In: Nachrichtenblatt des Deutschen Pflanzenschutzdienstes Bd. 59 (Nr. 7): 155–161.

SCHIMMELPFENNIG, S., J. ANTER and C. HEIDECKE (2018): Bewässerung in der Landwirtschaft - Tagungsband zur Fachtagung am 11./12.09.2017 in Suderburg. Thünen Working Paper No. 85.

SCHLENKER, W. and D. B. LOBELL (2010): Robust negative impacts of climate change on African agriculture. In: Environmental Research Letters 5 (1): 14010.

SCHLENKER, W. and M. J. ROBERTS (2006): Nonlinear Effects of Weather on Corn Yields*. In: Review of Agricultural Economics 28 (3): 391–398.

SCHLENKER, W. and M. J. ROBERTS (2009): Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. In: Proceedings of the National Academy of Sciences of the United States of America 106 (37): 15594–15598.

SHARMA, S., S. SHARMA and A. ATHAIYA (2020): ACTIVATION FUNCTIONS IN NEURAL NETWORKS. In: International Journal of Engineering Applied Sciences and Technology 04 (12): 310–316.

SUN, J., L. DI, Z. SUN, Y. SHEN and Z. LAI (2019): County-Level Soybean Yield Prediction Using Deep CNN-LSTM Model. In: Sensors (Basel, Switzerland) 19 (20).

TEIGEN, L. D. and M. THOMAS (1995): Weather and Yield, 1950-94: Relationships, Distributions, and Data.

TORETI, A., A. BELWARD, I. PEREZ-DOMINGUEZ, G. NAUMANN, J. LUTERBACHER, O. CRONIE, L. SEGUINI, G. MANFRON, R. LOPEZ-LOZANO, B. BARUTH, M. BERG, F. DENTENER, A. CEGLAR, T. CHATZOPOULOS and M. ZAMPIERI (2019): The Exceptional 2018 European Water Seesaw Calls for Action on Adaptation. In: Earth's Future 7 (6): 652–663.

TRNKA, M., J. E. OLESEN, K. C. KERSEBAUM, R. P. RÖTTER, R. BRÁZDIL, J. EITZINGER, S. JANSEN, A. O. SKJELVÅG, P. PELTONEN-SAINIO, P. HLAVINKA, J. BALEK, H. ECKERSTEN, A. GOBIN, V. VUCˇETIĆ, A. DALLA MARTA, S. ORLANDINI, V. ALEXANDROV, D. SEMERÁDOVÁ, P. ŠTEˇPÁNEK, E. SVOBODOVÁ and K. RAJDL (2016): Changing regional weather-crop yield relationships across Europe between 1901 and 2012. In: Climate Research 70 (2): 195–214.

TURVEY, C. G. (2001): Weather Derivatives for Specific Event Risks in Agriculture. In: Review of Agricultural Economics 23 (2): 333–351.

VAN KLOMPENBURG, T., A. KASSAHUN and C. CATAL (2020): Crop yield prediction using machine learning: A systematic literature review. In: Computers and Electronics in Agriculture 177: 105709.

VEDENOV, D. V. and B. J. BARNETT (2004): Efficiency of Weather Derivatives as Primary Crop Insurance Instruments.

VROEGE, W., J. BUCHELI, T. DALHAUS, M. HIRSCHI and R. FINGER (2021): Insuring crops from space: the potential of satellite-retrieved soil moisture to reduce farmers' drought risk exposure. In: European Review of Agricultural Economics.

WEBBER, H., G. LISCHEID, M. SOMMER, R. FINGER, C. NENDEL, T. GAISER and F. EWERT (2020): No perfect storm for crop yield failure in Germany. In: Environmental Research Letters 15 (10): 104012.

WHEELER, T. and J. von BRAUN (2013): Climate change impacts on global food security. In: Science (New York, N.Y.) 341 (6145): 508–513.

WOLANIN, A., G. MATEO-GARCÍA, G. CAMPS-VALLS, L. GÓMEZ-CHOVA, M. MERONI, G. DUVEILLER, Y. LIANGZHI and L. GUANTER (2020): Estimating and understanding crop yields with explainable deep learning in the Indian Wheat Belt. In: Environmental Research Letters 15 (2): 24019.

WOODARD, J. D. and P. GARCIA (2008): Basis risk and weather hedging effectiveness. In: Agricultural Finance Review 68 (1): 99–117.