

Comparing a random forest based prediction of winter wheat yield to historical yield potential

Yannik E. Roell¹, Amélie Beucher¹, Per. G. Møller², Mette B. Greve¹, Mogens H. Greve¹

¹ Department of Agroecology, Aarhus University, Denmark

² Department of History, Southern Denmark University, Denmark

(yannik.roell@agro.au.dk)

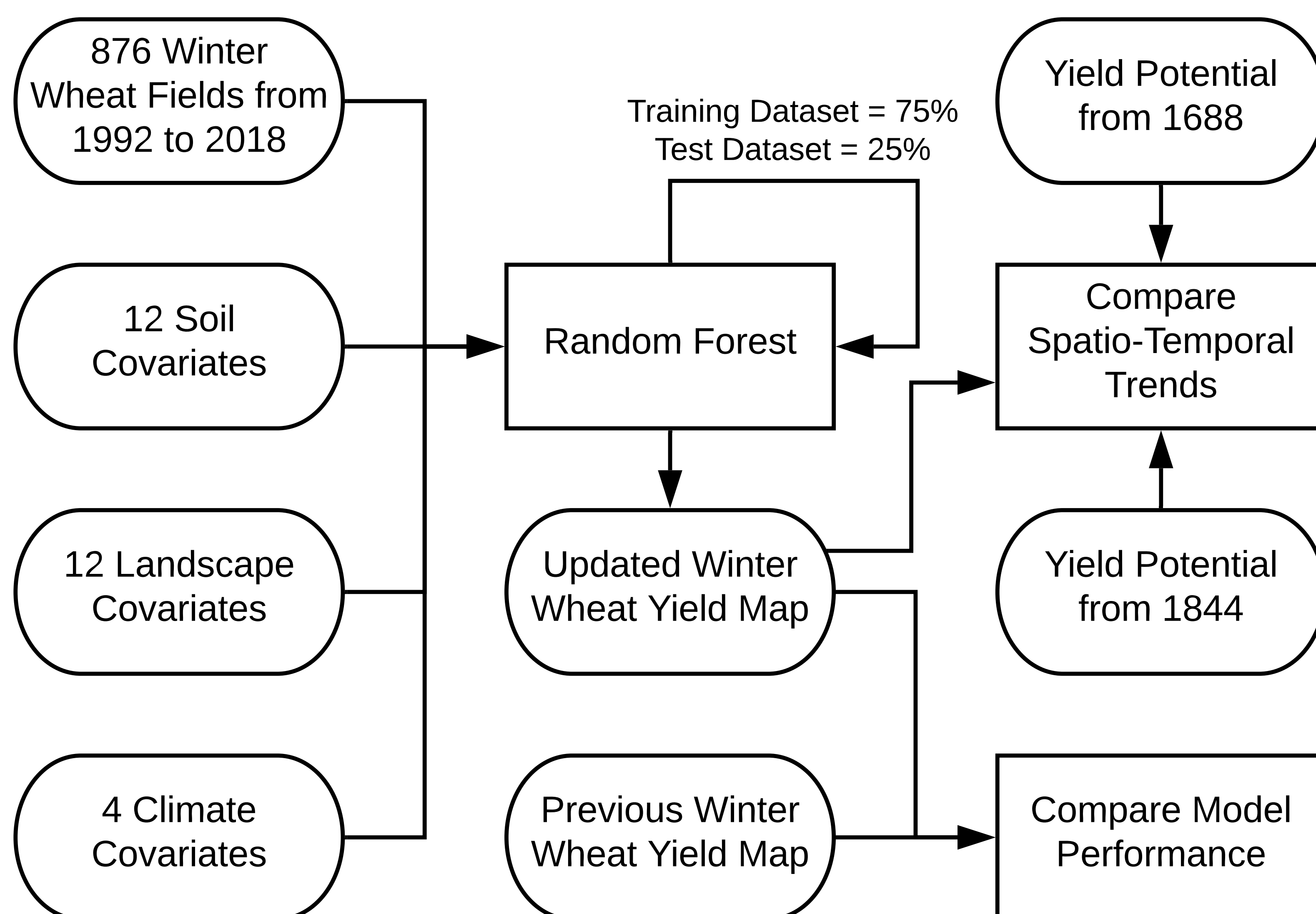


Introduction

- Winter wheat covers 23% of total cultivated land in Denmark.
- Denmark has detailed yield potential for the entire country as early as 17th century.
- Previous effort to model winter wheat for Denmark was based on a statistical model only utilizing three soil properties.

$$\text{Yield} = (\% \text{ Clay} + \% \text{ Silt}) + R_H - (\% \text{ Clay} + \% \text{ Silt}) \times R_H$$
 - where R_H is $(\% \text{ Clay} + \% \text{ Silt}) / \% \text{ Organic Carbon}$.
- This study: 1) updates the current national winter wheat yield map of Denmark and 2) compares current yield to historical yield potential.

Methodology



Random Forest Results

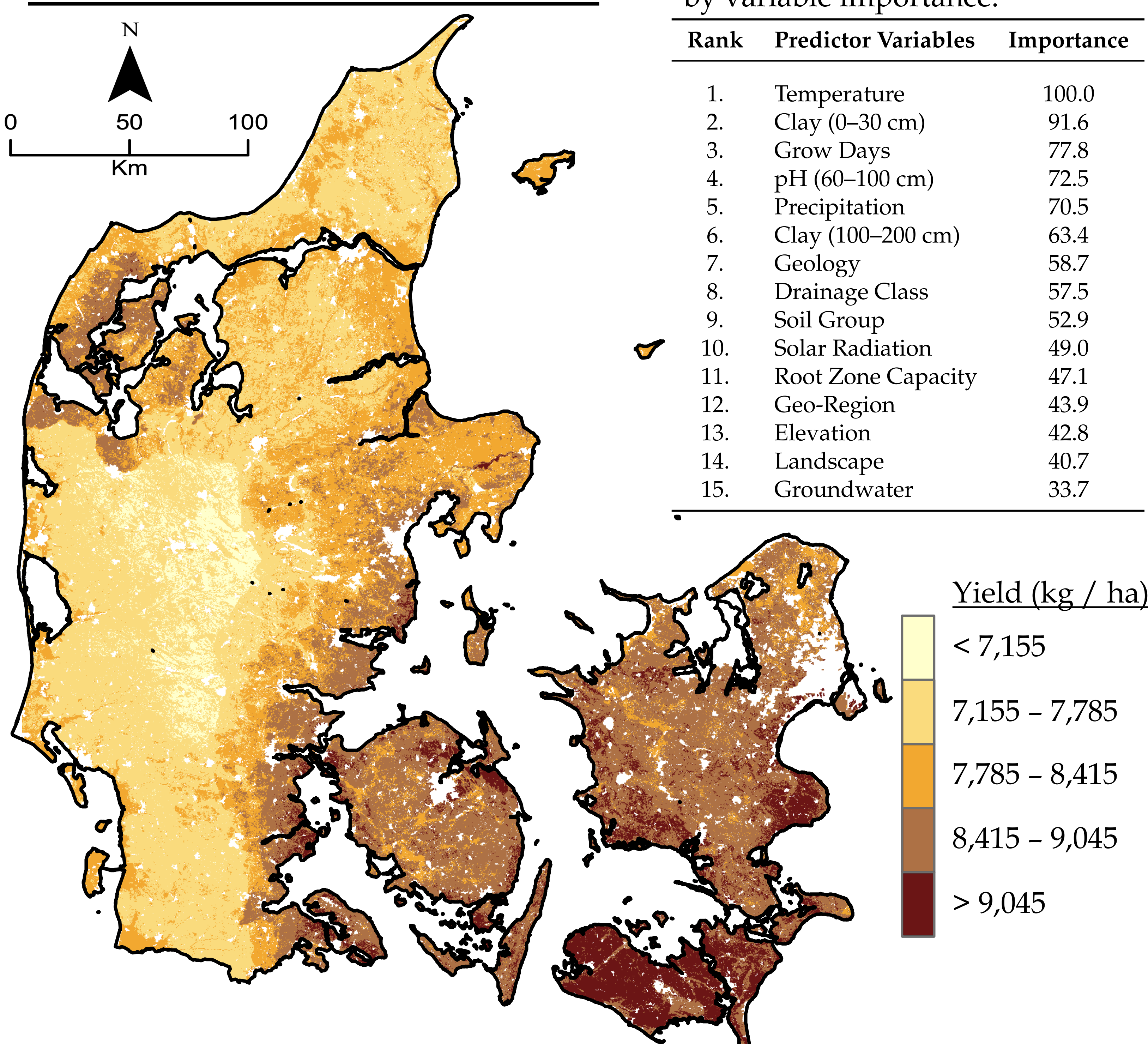


Figure 1: Yield map of winter wheat produced from random forest model.

Model Comparison Results

Figure 2: Comparison of the soil-only and random forest models using a random test dataset (n = 219; 25% of points).

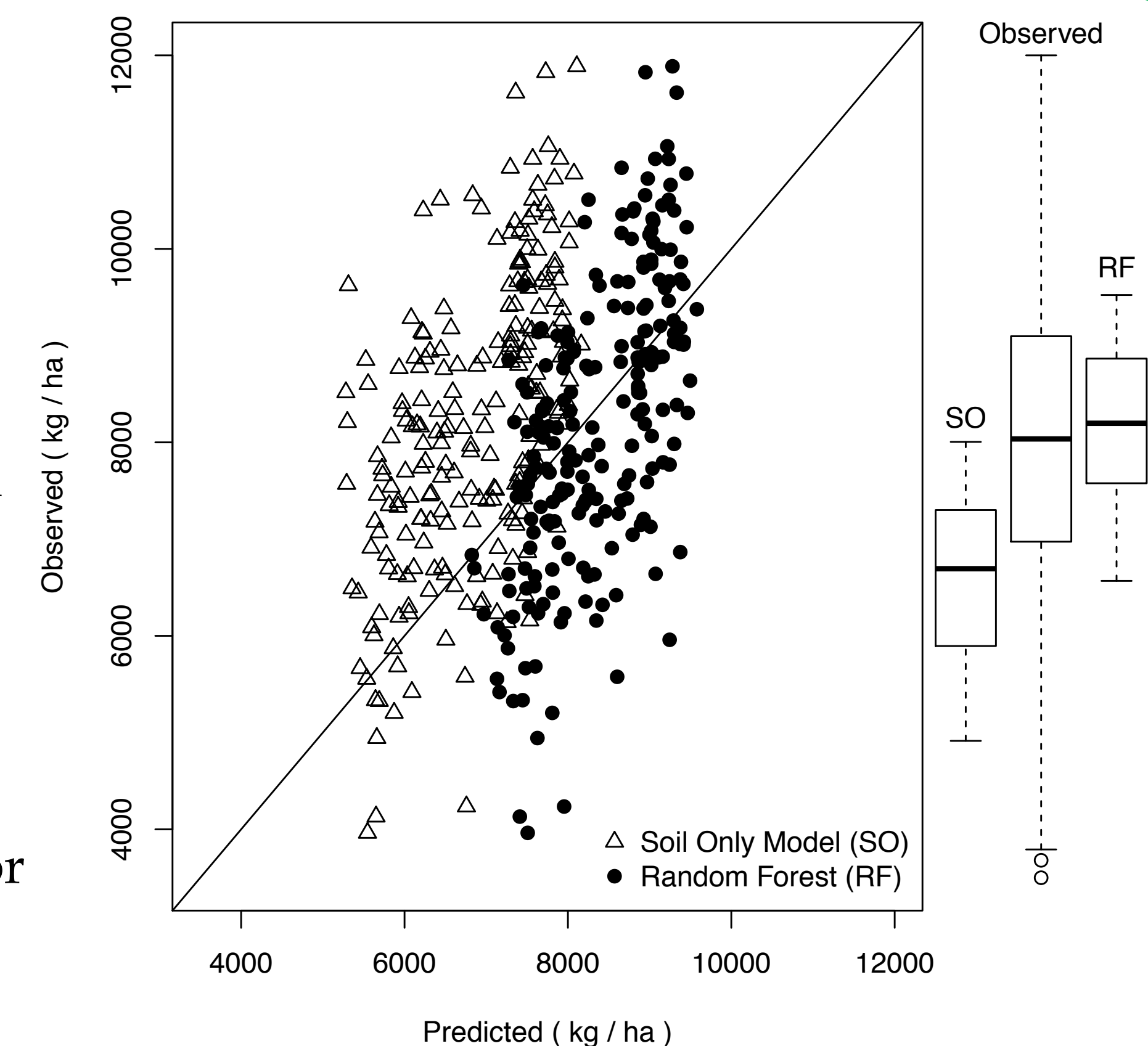


Table 2: Model performance measures for soil-only and random forest models.

Model		MSE	RMSE	CCC	R ²
Soil Only Model	Mean (SD)	3.93 × 10 ⁶ (2.3 × 10 ⁶)	1982 (60)	0.21 (0.03)	0.20 (0.04)
	Range	3.41 – 4.51 × 10 ⁶	1850 – 2120	0.14 – 0.28	0.11 – 0.31
Random Forest Model	Mean (SD)	1.58 × 10 ⁶ (1.49 × 10 ⁶)	1256 (60)	0.38 (0.03)	0.26 (0.05)
	Range	1.28 – 2.13 × 10 ⁶	1130 – 1460	0.31 – 0.45	0.18 – 0.37

MSE = mean squared error; RMSE = root mean squared error; CCC = Lin's concordance correlation coefficient; SD = standard deviation.

Spatio-Temporal Comparison Results

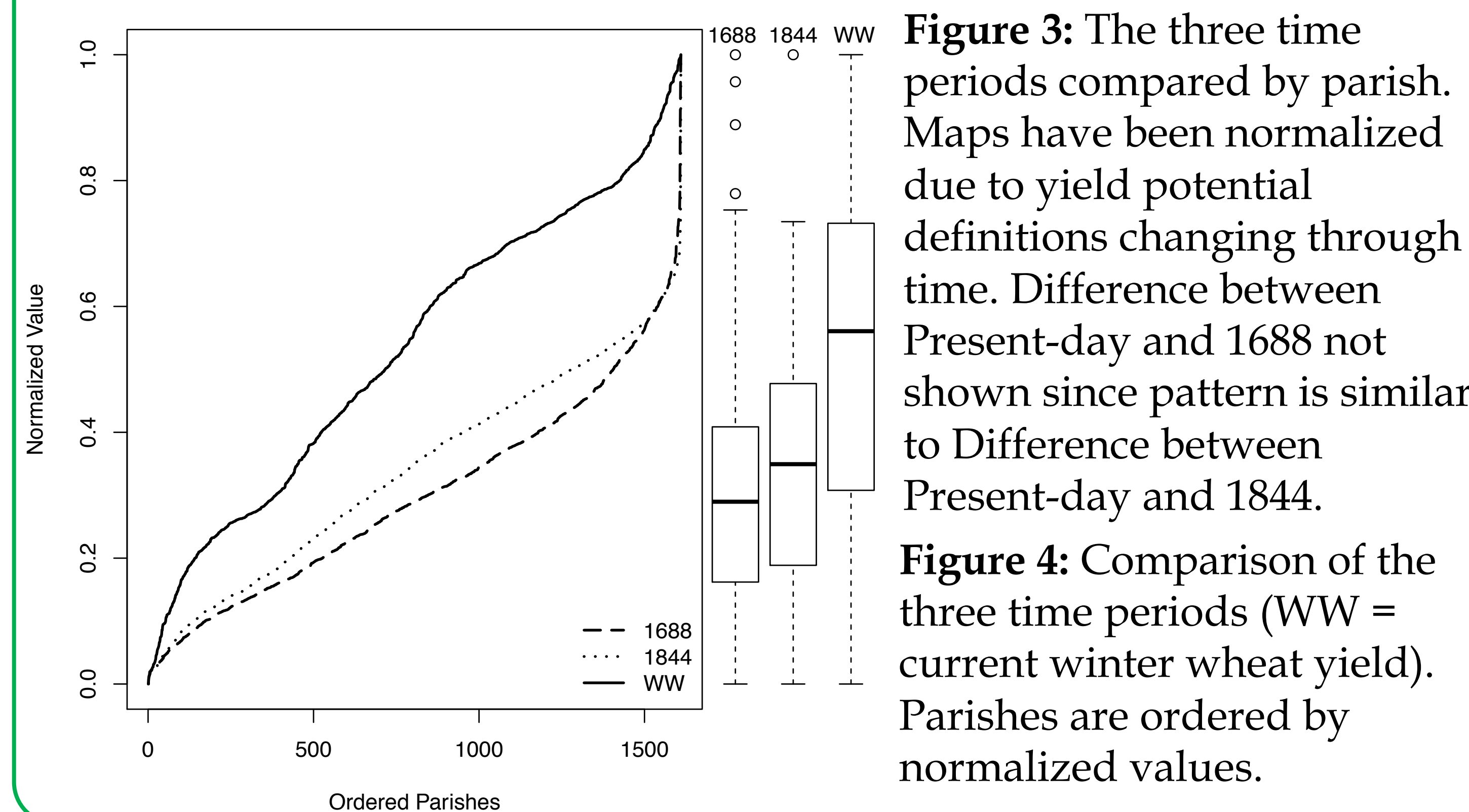
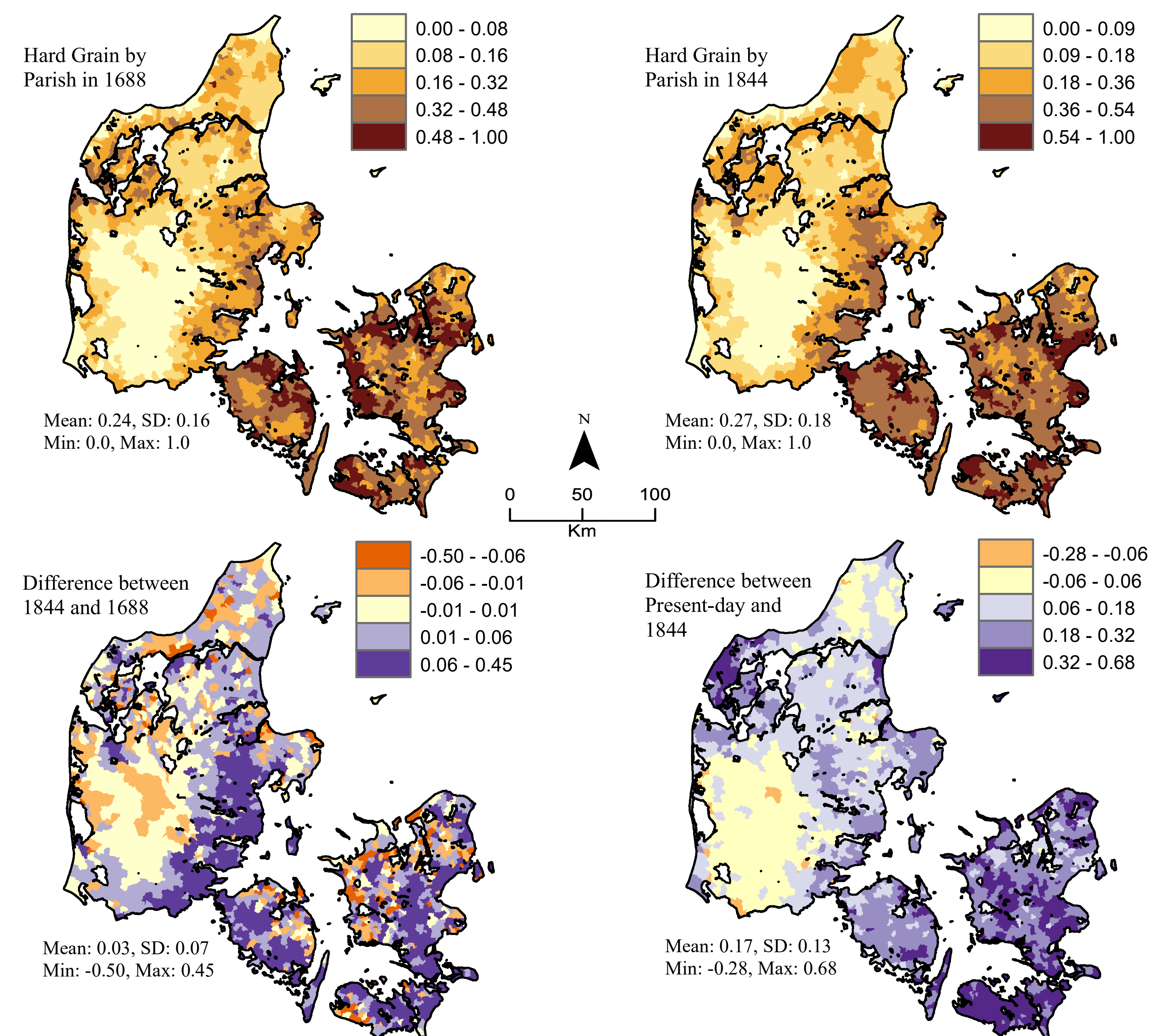


Figure 3: The three time periods compared by parish. Maps have been normalized due to yield potential definitions changing through time. Difference between Present-day and 1688 not shown since pattern is similar to Difference between Present-day and 1844.

Figure 4: Comparison of the three time periods (WW = current winter wheat yield). Parishes are ordered by normalized values.

Conclusions

- Random forest model performed better than the model based only on soil in all performance measures.
- Historical time periods are characterized by numerous low yield potential areas and few highly productive areas; current yield is evenly distributed between low and high yields.

Article

Comparing a Random Forest Based Prediction of Winter Wheat Yield to Historical Yield Potential

Yannik E. Roell ^{1,*} , Amélie Beucher ¹ , Per G. Møller ², Mette B. Greve ¹ and Mogens H. Greve ¹

¹ Department of Agroecology, Aarhus University, Blichers Allé 20, 8830 Tjele, Denmark; amelie.beucher@agro.au.dk (A.B.); metteb.greve@agro.au.dk (M.B.G.); greve@agro.au.dk (M.H.G.)

² Department of History, Southern Denmark University, Campusvej 55, 5230 Odense, Denmark; pgm@sdu.dk

* Correspondence: yannik.roell@gmail.com

Received: 19 December 2019; Accepted: 10 March 2020; Published: 14 March 2020



Abstract: Predicting wheat yield is crucial due to the importance of wheat across the world. When modeling yield, the difference between potential and actual yield consistently changes because of advances in technology. Considering historical yield potential would help determine spatiotemporal trends in agricultural development. Comparing current and historical yields in Denmark is possible because yield potential has been documented throughout history. However, the current national winter wheat yield map solely uses soil properties within the model. The aim of this study was to generate a new Danish winter wheat yield map and compare the results to historical yield potential. Utilizing random forest with soil, climate, and topography variables, a winter wheat yield map was generated from 876 field trials carried out from 1992 to 2018. The random forest model performed better than the model based only on soil. The updated national yield map was then compared to yield potential maps from 1688 and 1844. While historical time periods are characterized by numerous low yield potential areas and few highly productive areas, current yield is evenly distributed between low and high yields. Advances in technology and farm practices have exceeded historical yield predictions, mainly due to the use of fertilizer, irrigation, and drainage. Thus, modeling yield projections could be unreliable in the future as technology progresses.

Keywords: yield modeling; machine learning; climate; yield potential; historical yield

1. Introduction

Developing accurate national yield models is challenging due to variations in growth conditions across large areas and changes through time. A range of models have been created to simulate crop growth because of the importance of predicting yield accurately (e.g., CERES-Wheat: [1]; WOFOST: [2]; Sirius: [3]). Correctly modeling yield is necessary for an array of crops, but a large number of studies have focused on wheat. Wheat is a staple food globally, corresponding to approximately 30% of total world cereal production, and provides a large portion of both daily calories and protein for people across the world [4]. Wheat is grown in nearly all parts of Europe [5], spanning a wide range of soil types and climate zones, meaning the models are required to be diverse and robust.

The importance of wheat has driven research to determine what affects yield potential [6–10]. Many decisions made by farmers alter yields (crop variety, planting date, irrigation, etc.); however, many factors affecting yield potential are independent from the farmer's decisions. Three major components affecting yield are: soil, climate, and topography. Soil properties, such as structure and clay percentage, either hinder or promote growth [11]. Fluctuations in temperature and precipitation have been found to be key reasons for uncertainty in yield predictions [12]. Topography can impact yield by affecting water flow, energy supply, and accumulation of soil and nutrients in an area [13,14].

Incorporating aspects of soil, climate, and topography in modeling yield is difficult due to the complexity of these environmental factors.

Recently, many studies have used machine learning (ML) techniques, instead of statistical methods, as a way to predict yields. ML algorithms allow for a high degree of complexity and are capable of learning from large datasets. While some ML techniques are viewed as a black box, many linear and nonlinear relationships can be discovered using a sole ML algorithm that requires different statistical models [15,16]. Another benefit of ML is the ability to integrate different types of data easily. Data fusion has been a key focus for improving wheat yield predictions. National yield models could be improved by coupling data, such as climate and sensor data (i.e. from satellite images and proximal soil sensing), with ML algorithms to predict yield [17,18]. Most models make predictions of current yield, but as technology advances, current yield models will be unreliable because the yield potential will continue to change.

The difference between yield potential and actual yield indicates that limiting factors exist in the field while a crop is growing [19]. Yield potential is defined as a crop that has adapted to the environment and is not limited in terms of water and nutrients [20]. Advances in technology will increase both potential and actual yields of a crop. To assess the changes in yield, studies typically look at historical yield within the 20th century [21,22]. Denmark, on the other hand, has detailed yield potential for the entire country from as early as the 17th century [23]. *King Christian V's Great Danish Land Register of 1688* and *The Great Danish Land Register of 1844* provided valuable information to determine the natural growth potential before advances in mechanical farm technology were wide spread. These land registers are comprehensive for the entire country, and include land evaluations. Both of these land registers used hard grain values to evaluate the land. Hard grain, a measure used to determine how much was payable to landlords, was equivalent to a certain amount of a product (rye, barley, wheat, honey, geese, etc.) and was based on area and soil quality [23,24].

Denmark has a large portion of the country dedicated to arable land (over 60% [25]), and has consistently been highly farmed throughout history. The high level of farming through the years provides a unique opportunity to compare historical yield potential with current yields. The changes may indicate how advances in technology and farming practices have driven agricultural production. Winter wheat covers 23% of the total cultivated land in Denmark [25], making winter wheat an important crop for the country. However, the previous effort to model winter wheat for Denmark was based on a statistical model only utilizing soil properties [26]. The following model was developed to predict yield from soil properties:

$$Y_{opt} = 4.78 + 6.67 \times (\% \text{ Clay} + \% \text{ Silt}) + 0.11 \times R_H - 0.21 \times (\% \text{ Clay} + \% \text{ Silt}) \times R_H \quad (1)$$

where Y_{opt} is predicted yield in t ha^{-1} and R_H is $(\% \text{ Clay} + \% \text{ Silt}) / \% \text{ Organic Carbon}$. The model includes clay, silt, and organic carbon of the topsoil (0–30 cm) only.

This study has two main objectives: (1) to update the current national winter wheat yield map of Denmark, and (2) to compare current yield to historical yield potential. The first objective uses an array of soil, climate, and topography variables with random forest to generate a more accurate yield map. The updated yield map is compared to the existing yield map to ensure the addition of climate and topography have increased yield prediction accuracy. The second objective will use historical yield potential maps from 1688 and 1844 for comparisons with the newly developed yield map. Comparing yield through time will help determine how advances in technology have driven agriculture in Denmark.

2. Materials and Methods

2.1. Study Area

Denmark is a Northern European country lying between 54.3–57.5° N and 8.0–15.2° E with a temperate climate. The mean annual temperature is 9 °C with summer days averaging to 17 °C and

winter days averaging to 1 °C. Mean annual precipitation is 770 mm, with the west receiving more rain than the east [27]. Multiple glaciations during the Weichselian geological period have resulted in highly variable soil distribution and formation [28]. Soils are mainly loamy Luvisols in eastern and central regions and sandy Podzols in western regions [29]. Moraine landscapes dominate the east and west, with the west being glaciofluvial flood plains and older eroded moraine landscapes. Raised seabeds dominate the north. The country covers 43,000 km² of area, and due to many glacial advances, topography is relatively flat with a mean elevation of 32 m and a maximum elevation of 171 m.

2.2. Winter Wheat Fields

Winter wheat (*Triticum aestivum* L., spp. *vulgare*) yields were acquired from the Danish National Field Trials managed by the Danish Agricultural Advisory Service [30]. The data from 876 fields from 1992 to 2018 were used in this study (Figure 1). Bornholm was excluded due to the lack of fields on the island. Each field follows an experimental design that uses 4 or 5 randomized blocks (30 m² plot sizes) with every block receiving a different level of mineral nitrogen fertilizer. The maximum yield (in kg ha⁻¹) per field was used in this study. The average mineral nitrogen fertilizer applied was 202 kg N ha⁻¹.

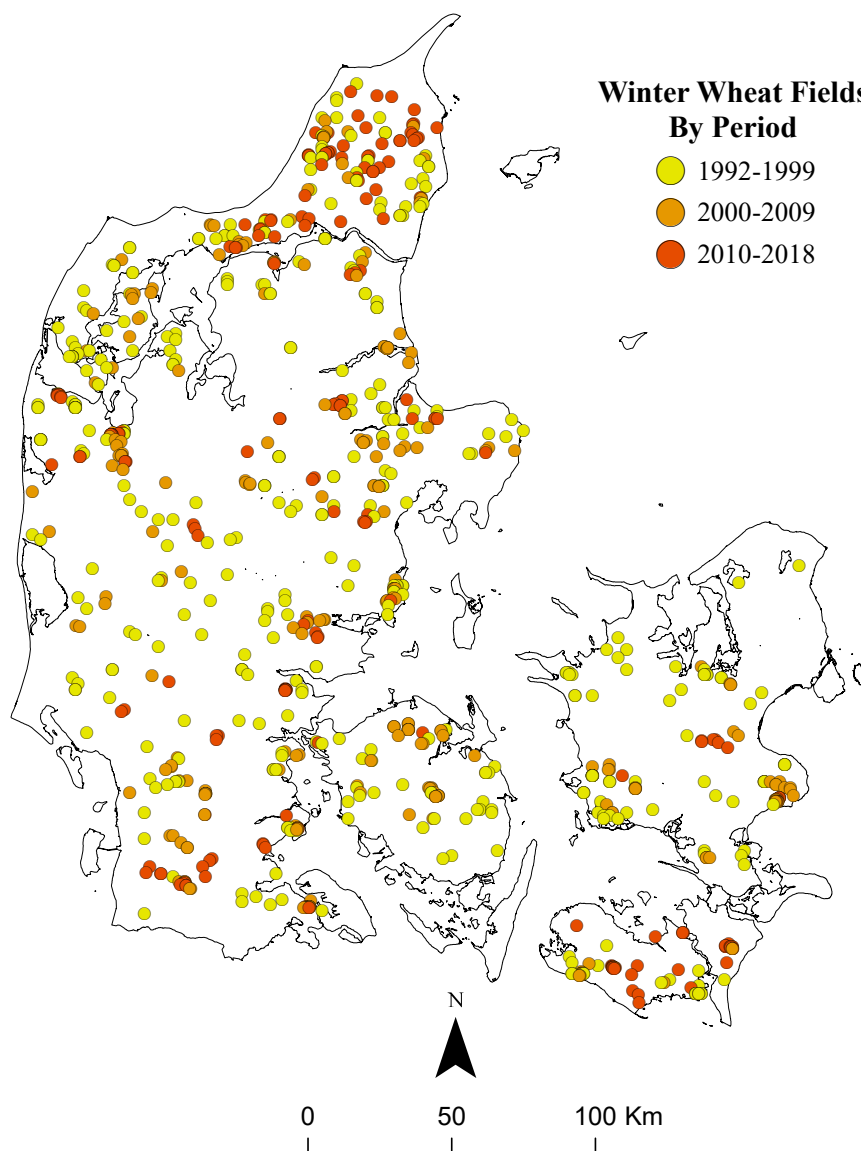


Figure 1. Locations of winter wheat fields (n = 876) from 1992 to 2018.

Fields covered all climate regions and soil types in Denmark, as different locations were selected each year. An average of 32 fields was selected each year (range between 14 and 60). Yield increased gradually from 1992 to 2018 (Figure 2A). For every year, average yield increases by 70 kg ha^{-1} (linear regression: $p\text{-value} < 0.001$). Thus, the trend caused by differences between years was removed (Figure 2B). This increase could be attributed to many factors (advances in farming technology, better management techniques, higher annual temperature, etc.). The trend was removed by subtracting a correction factor from each yield value. The correction factor is calculated by subtracting the global average from the year's average, resulting in different correction factors for each year.

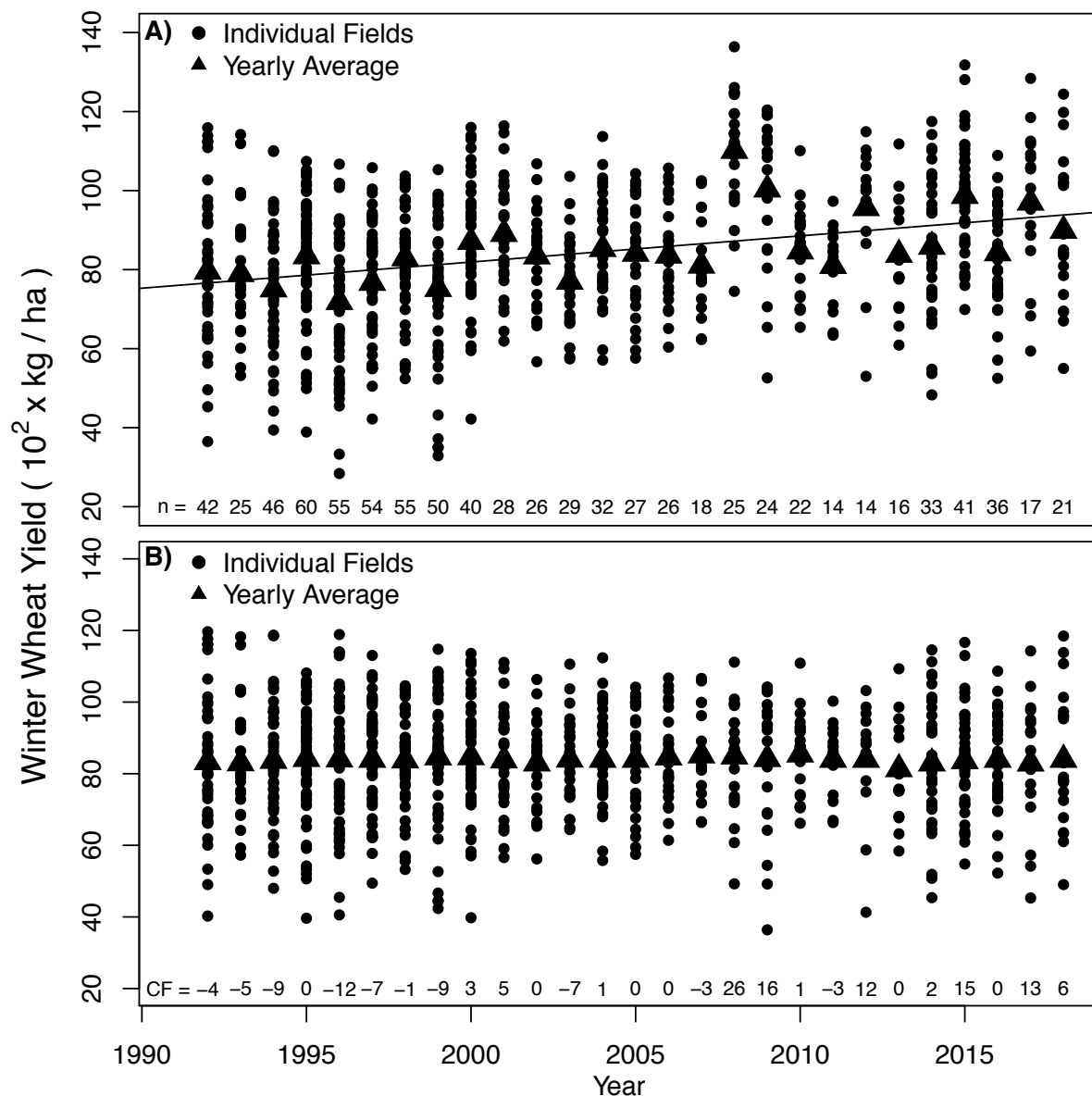


Figure 2. (A) Yield from winter wheat fields exhibiting a positive increase through the years and (B) yield detrended by subtracting a correction factor (CF; yearly average – global average) from individual fields to remove temporal trend.

2.3. Yield Modeling

Random forest (RF) is used to model winter wheat yields from field trials. This method is a supervised decision tree algorithm that bags un-pruned trees by using randomly selected covariates at each split [31]. The RF method grows a large number of trees from a set of independent observations.

Randomness is used to select a variable at each split in the tree, and this variable is from a random subset of all predictor variables. The RF model only generates the mean value for each tree, but using quantile regression forest [32], the distribution of values predicted for each tree can be utilized. The difference in values from quantiles (0.1 and 0.9) is used as the prediction interval for the model. Both methods are implemented from the *ranger* package in R [33].

The RF model was generated 100 times on different training sets to determine the model performance on different input data. Winter wheat fields are divided into training and test datasets. The training dataset represents 75% of the original data ($n = 657$) and was used to build the model; the test dataset represents the remaining 25% of the original data ($n = 219$) and was used to assess model performance. The final winter wheat yield map was generated using all of the data. Variable importance was calculated by averaging variable importance for all 100 model runs. The method used for calculating variable importance within the RF model was the permutation method.

The RF model was compared to the existing winter wheat yield map of Denmark [26]. The previous method used the same source to attain winter wheat yield values as the RF model. Values from the winter wheat yield map based only on soil data are extracted at each field location to determine whether the RF model improved yield predictions. To ensure that the correction factor for a year did not create a bias towards the RF model, uncorrected and corrected values were compared to both the RF and soil-only models. Model performance was assessed using: mean squared error (MSE), root mean squared error (RMSE), Lin's concordance correlation coefficient (CCC), and coefficient of determination (R^2).

2.4. Predictor Variables

Covariates are needed to build the RF model to find the association between the environment and yield. Covariates are divided into three groups (Table 1): soil (12 covariates), topography (12 covariates), and climate (4 covariates). Covariates were chosen based on expert knowledge and previous studies indicating the predictor variables affecting crop growth. All covariates were generated or resampled to have a 30.4-m spatial resolution. Danish variables created for digital soil mapping typically have a 30.4-m resolution because the digital elevation model (DEM) was derived from 1.6-m resolution national LiDAR data and resampled to 30.4-m resolution [34].

Soil covariates were acquired from previous studies. Soil covariates were divided into two types: continuous and categorical. The continuous soil covariates are: organic carbon percentage (depths 0–30 and 100–200 cm) [35], clay percentage (depths 0–30 and 100–200 cm) [36], drainage class [37], depth to groundwater (60–100 cm) [38], pH [39], and winter wheat root zone capacity (based on a previously developed pedotransfer function) [24]. The categorical soil covariates are: soil group [34], geology [40], geo-region [36], and landscape [24]. Topography covariates are comprised of elevation and 11 derivatives based on the DEM [34].

Climate variables are based on temperature, solar radiation, and precipitation. Daily mean temperature in °C at 2 m above the ground and global solar radiation in MJ/m² were downloaded from Aarhus University's meteorological database [41]. Downloaded climate data were from 1985 to 2013 at a resolution of 40 km. Kriging was performed on 40-km grids to generate maps of: annual number of growing days above 10 °C, average annual temperature, and global solar radiation during the growing season (between 1 April and 31 October). The solar radiation map was combined with a DEM solar radiation map, generated from the Area Solar Radiation tool in ArcMap, to create a topographically corrected solar radiation map. The precipitation map was derived from average total rainfall during the growing seasons between 1961 and 1990 [42].

Table 1. Predictors used for random forest model of winter wheat yield.

Predictor Variables	Ranges/ Classes	Description	Source
<i>Soil</i>			
Carbon (0–30 cm)	0–55.86	Percent organic carbon in soil between 0 and 30 cm	[35]
Carbon (100–200 cm)	0–22.22	Percent organic carbon in soil between 100 and 200 cm	[35]
Clay (0–30 cm)	0–51.20	Percent clay in soil between 0 and 30 cm	[36]
Clay (100–200 cm)	0–51.96	Percent clay in soil between 100 and 200 cm	[36]
Drainage Class	1.10–4.61	Drainage classes between very poorly drained and very well drained soils	[37]
Soil Group	10 classes	Soil class map	[34]
Geology	10 classes	Geological map	[40]
Geo-Region	7 classes	Geographical regions map	[36]
Groundwater	0–143.1	Depth to groundwater in meters	[38]
Landscape	11 classes	Landform types	[24]
pH (60–100 cm)	2.37–9.62	pH in soil between 60 and 100 cm	[39]
Root Zone Capacity	11.2–1581.7	Weighted sum of winter wheat rooting depths using 5 depths in millimeters	[24]
<i>Topography</i>			
Elevation	0–169.9	LiDAR produced elevation of land surface	[34] ¹
Curvature (Plan)	−1.60–2.51	Curvature perpendicular to the maximum slope	
Curvature (Profile)	−3.14–3.36	Curvature parallel to the maximum slope	
Horizontal Distance to Channel Network	0–3223	Horizontal distance to the nearest waterbody	
Mid-Slope Position	0–1	Relative vertical distance to mid-slope position	
Multi-Resolution Valley Bottom Flatness	0.01–10.62	Calculates depositional areas to identify flat valley bottoms	
Relative Slope Position	0–1	Slope position and the relative position between valley and ridge	
SAGA Wetness Index	7.30–18.90	Topographic wetness index with modified catchment area	
Slope	0–90	Slope gradient in degrees	
Slope to Channel Network	0–47.63	Slope to the nearest waterbody	
Valley Depth	0–86.04	Relative height difference to adjacent channel	
Vertical Distance to Channel Network	0–111.45	Vertical distance to nearest waterbody	
<i>Climate</i>			
Grow Days	146.8–157.8	Number of days above 10 °C during the year	[41] ²
Precipitation	281.9–481.4	Precipitation between April and October in mm	[42]
Solar Radiation	1802–3422	Global solar radiation between April and October in MJ/m ²	[41] ²
Temperature	7.90–8.92	Average yearly temperature in °C	[41] ²

¹ Topography variables are derivatives of the digital elevation model [34] and were generated in SAGA.

² Raw data from AGRO Climate Database [41]; kriging was used on points to obtain the map.

2.5. Historical Comparison

Historical yield potential values are used to assess temporal changes of production through Denmark. The 1688 ([23]) and 1844 ([43]) land registers have been mapped using previously digitized data [44]. The proper data corrections were made to ensure both time periods aligned spatially. Hard grain ha^{−1} values from 1688 and 1844 were assigned to each parish from 1820. Land in the registers was evaluated based on yield potential using a unit of measure called barrels of hard grain, but different definitions were used to quantify hard grain in 1688 than in 1844. Thus, values were normalized from 0 to 1 to determine relative differences. In both definitions, hard grain was based on field measurements and evaluations which could be expressed in terms of production.

The three time periods (1688, 1844, and current) are compared to determine the spatial trend in agricultural production. The winter wheat yield values were averaged for each parish and normalized. The first two time periods assess yield potential, and the current values determine actual yield. While the comparison is based on relative values, the distribution of values for each time period will indicate how national farming has progressed. The comparisons will illuminate areas that have always been viewed as highly productive and areas that have recently been developed due to advances in farming.

3. Results and Discussion

3.1. Winter Wheat Yield Predictions

3.1.1. Random Forest Predictor Variables

The predictor variable importance indicates that using climate is crucial when predicting winter wheat yield (Table 2). All four climate variables were within the top 10 predictor variables and significantly correlated with yield. Organic carbon at a depth of 0–30 cm was significant and negatively correlated with yield, but had low importance. Organic carbon was one of the three properties used in the soil-only model, but neither organic carbon variable (depths 0–30 and 100–200 cm) was within the top 15 predictors. Clay at a depth of 0–30 cm was used in the soil-only model and was an important predictor for yield; however, clay at a depth of 100–200 cm was also crucial in the RF model, but was not utilized in the soil-only model. Winter wheat roots have been shown to reach 2 m depth [45], indicating subsoil properties are necessary for predicting winter wheat yield. This shows that the soil-only model might be too simple to accurately predict yield values.

Table 2. Top 15 predictors ordered by importance in the random forest model using permutation as the method to determine variable importance. Correlation is indicated by a + for positive correlation, - for negative correlation, and C for categorical variable. A bold + or - indicates whether predictors are significantly correlated with winter wheat yield (linear regression: p -value < 0.05).

Rank	Predictor Variables	Importance	Correlation
1.	Temperature	100.0	+
2.	Clay (0–30 cm)	91.6	+
3.	Grow Days	77.8	+
4.	pH (60–100 cm)	72.5	+
5.	Precipitation	70.5	-
6.	Clay (100–200 cm)	63.4	+
7.	Geology	58.7	C
8.	Drainage Class	57.5	+
9.	Soil Group	52.9	C
10.	Solar Radiation	49.0	+
11.	Root Zone Capacity	47.1	-
12.	Geo-Region	43.9	C
13.	Elevation	42.8	-
14.	Landscape	40.7	C
15.	Groundwater	33.7	-

Topography variables did not have a large influence on yield predictions. Elevation was the only variable within the top 15 variables. While some of the topography variables were significantly correlated with yield, the overall importance of these variables was low. Contrary to our results, studies have found that topographic position has a crucial effect on wheat yield. While topography explained a lower amount of variation in yield than soil properties, Kravchenko and Bullock [46] indicated elevation was the most influential topographic variable on yield. However, due to the Danish landscape being relatively flat, the addition of topography variables does not play a large role in yield predictions. With soil being an important factor for predicting yield, and topography being a large component in modeling soil, topography is indirectly associated with yield. All of the individual soil properties used in this study have incorporated topography when being modeled and mapped. Thus, adding topography variables alone is likely redundant and seems to have only made the RF model more complicated and computationally expensive.

The addition of climate variables in modeling yield is an important aspect many researchers are focusing on, due to shifts in climate patterns. In varying climate simulations, yield was predicted to decrease with an increase in temperature variability, while changes in yield were mediated by soil type

when precipitation was varied [47]. In yield forecasting, models incorporating seasonal temperature and precipitation were able to reliably predict yield variability for 25–38% of the globally harvested area before harvesting [48]. Kristensen et al. [49] stated winter wheat yield in Denmark will likely decrease with the projected climate change and Webber et al. [50] concluded the same results but for all of Europe. Yield has been predicted to decrease by 550 kg ha⁻¹ from 1985 to 2040 with varying degrees of yield reduction depending on the location and soil type. Using the same field trial data source, Ørum et al. [51] found many aspects of climate were important for predicting winter wheat yield in Denmark. Average temperature and precipitation, during certain months on varying soil types, impact plant growth significantly [52]. Besides climate variables, atmospheric carbon dioxide (CO₂) should be included in modeling winter wheat yield because the increase in CO₂ and the resulting interaction with environmental variables are expected to increase yield [53]. Broberg et al. [54] elevated CO₂ levels from 372 to 605 ppm and found winter wheat yield increased by 26%. While the seasonal climate during the growing period does not change drastically in Denmark, climate variables and CO₂ need to be included in future yield models.

Table 3. Model performance measures for soil-only and random forest models.

Model		MSE	RMSE	CCC	R ²
Soil Only Model	Mean (SD)	3.93 × 10 ⁶ (2.3 × 10 ⁶)	1982 (60)	0.21 (0.03)	0.20 (0.04)
	Range	3.41–4.51 × 10 ⁶	1850–2120	0.14–0.28	0.11–0.31
Random Forest Model	Mean (SD)	1.58 × 10 ⁶ (1.49 × 10 ⁶)	1256 (60)	0.38 (0.03)	0.26 (0.05)
	Range	1.28–2.13 × 10 ⁶	1130–1460	0.31–0.45	0.18–0.37

MSE = mean squared error; RMSE = root mean squared error; CCC = Lin's concordance correlation coefficient; SD = standard deviation.

3.1.2. Model Comparison

The RF model used for predicting winter wheat yield performed better than the previous winter wheat yield model based only on soil properties in all model performance evaluations (Table 3). The correction factor applied to the winter wheat values did not bias the RF model, since the soil-only model performed better when comparing to corrected values to uncorrected values. The RF model's MSE was half the one of the soil-only model, and the RF model's RMSE was 726 kg ha⁻¹ lower than that of the soil-only model. Lin's CCC constitutes a measurement for both precision and accuracy [55], and a valuable way to measure agreement between two methods measuring the same variable. The coefficient was ranked poor for the soil-only model (0.21) and nearly ranked fair to good (0.38) for the RF model using the Fleiss interpretation [56]. The coefficient of determination (R²) is 0.06 higher for the RF model than for the soil-only model. This indicates the RF model explains 6% more variation than the previous model. Looking at model performances, while the R² did not increase drastically, the incorporation of climate and topography are beneficial in modeling winter wheat yield, as indicated by every other performance measurement.

Distributions of the predicted yields for both models do not account for extreme values, as seen in the raw data (Figure 3). The RF model did not perform well when calculating extreme values, but the predictions are centered around the observed data, whereas the soil-only model consistently predicts low yields. The ability to predict extreme values could be improved by incorporating farm management data. Integrating information on irrigation and the previous crops planted have shown to be significant factors that impact yield [51]. However, generating national maps of farm management would be time-consuming and expensive. The placement of field trials could be another reason why the model does not have a higher R² value and topography does not have a large contribution. The range of slope for agricultural land in Denmark is between 0 and 34°, while the range of slope for the field trials is between 0 and 12°. The fields change every year but the placement is not completely random, resulting in a biased sampling scheme that does not consider extreme soil and topography locations.

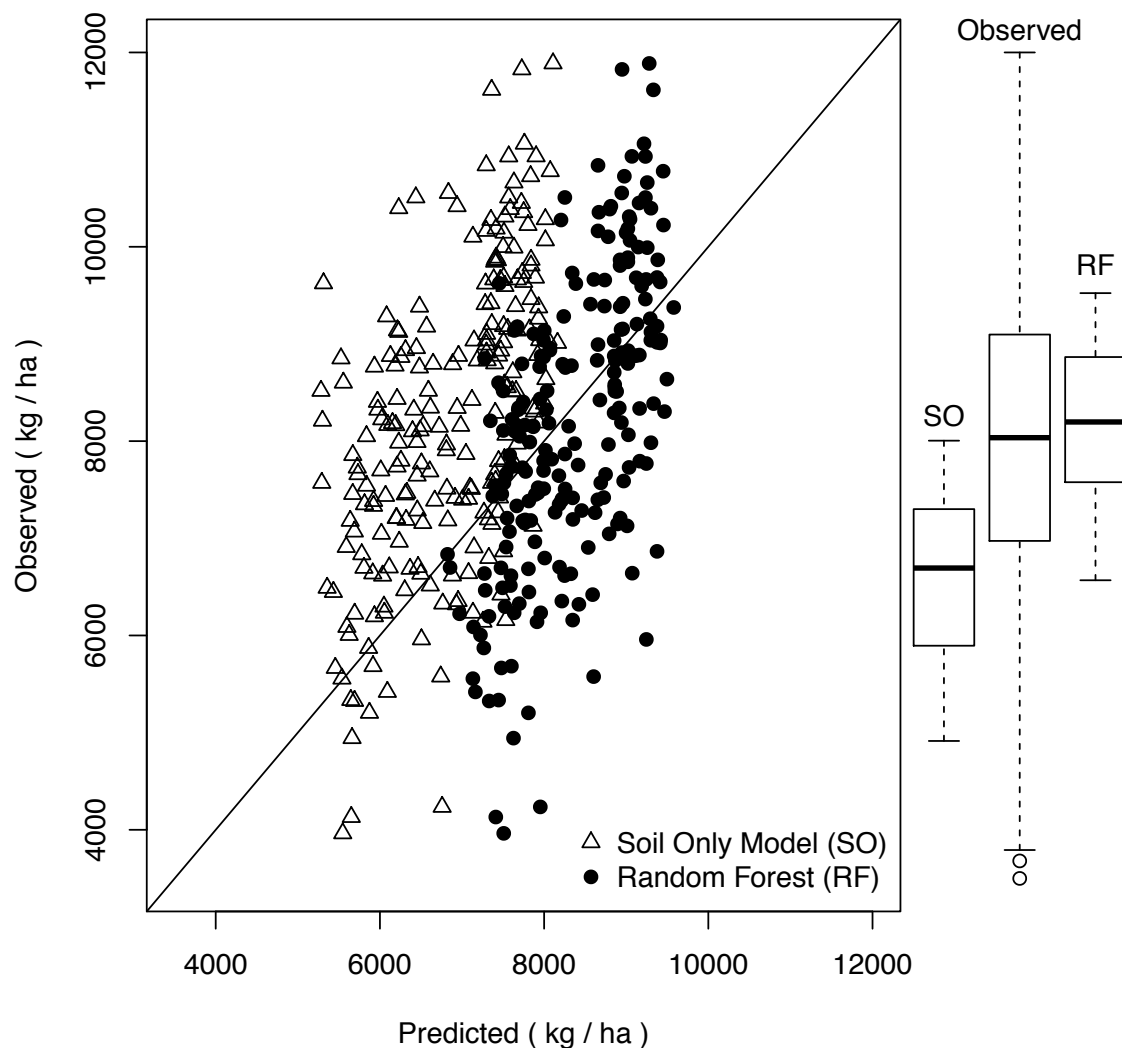


Figure 3. Comparison of the soil-only and random forest models using a random test dataset ($n = 219$; 25% of points).

Other studies have compared the use of ML algorithms to linear regressions in crop yield studies. Linear regression has out-performed neural networks when using only one input, but neural networks consistently out-performed linear regression when multiple inputs were used [57]. Random forest has been found to produce smaller RMSE values than linear regression when dealing with regional and global scale crop yield predictions [58]. In a study calculating solar radiation from both statistical models and neural networks, the neural networks provided the best estimates on the training data but performed poorly in a test site with different weather conditions. In contrast, the statistical models achieved adequate results in both the training and test sites [59]. The ability to find complex relationships between multiple inputs is the benefit of ML algorithms; however, these algorithms do fall short when applied to datasets with few inputs or when new data arises in test datasets.

The final map from the RF model along with prediction interval (difference in quantiles 0.1 and 0.9) is shown in Figure 4A,B. The pattern indicates that the best yield for winter wheat is in the east with loamy soils and the worst in the west with fluvial sands. This result is consistent with the soil fertility map from 1844 [60] and the existing yield map (Figure 4C). Areas with the highest and lowest yields also displayed the lowest prediction intervals. This indicates that these areas are typically stable and do not fluctuate as much as other areas. The stability in yield prediction is probably due to soil quality either being good or poor, resulting in consistent yield no matter the climate. The largest prediction intervals are associated with areas that have intermediate yield values. The large prediction

interval likely indicates that the yearly fluctuations are greatest in these areas, since the soil quality is average and not a reliable buffer. Thus, climate would have a larger influence on yield.

The two yield maps were normalized between 0 and 1 and subtracted from each other (Figure 4D). The maps were normalized because the soil-only model predicts negative values in certain areas. These negative values appear for locations where the organic carbon content is below 1% and much smaller than the clay and silt contents. Thus, the soil-only yield map needed to be normalized. The RF model predicted higher values in the east and lower values in the west compared to the previous model. Areas with average yield values are roughly in the same location in both models. The main differences are in the extreme values where the yield is either the highest or lowest. As seen in Figure 4C, the previous model generates a smoother map because only three soil properties were used in the model. The soil-only model originated from work that was focusing on the effects soil organic matter has on crop yield when optimal nitrogen was allocated [26], creating a model that did not incorporate environmental factors. The RF model has sharper boundaries because of variables that take into account glacial boundaries. The lowest values in the RF model from the middle of Jutland are indicative of the boundary between the young moraine and outwash plain landscape types [24]. The fine soil particles (clay and silt) are typically above 20% in the young moraine, resulting in a larger supply of plant available water, and below 10% in the outwash plain, resulting in a smaller supply of plant available water. These differences in soil genesis and soil texture lead to large yield differences in Jutland. The incorporation of variables explaining climate and geology into the model are crucial for depicting the actual pattern seen, rather than the simplified version only including soil.

3.2. Historical Comparisons

The three time periods are mapped by parish in Figure 5. Spatial patterns are similar for the first two time periods (1688 and 1844). Comparing the two historical periods, areas in the east had higher yield potentials in 1844 than 1688. Both time periods were similar in the west, and no clear pattern is seen in the north regarding which time period had a higher yield potential. The differences are similar between current yield and both historical yield potentials. The differences indicate that the changes between 1844 and 1688 are minor compared to the drastic changes in current yield. However, areas that have been the least productive are constantly the least productive. Although advances in technology have made the western part of Denmark more productive, the soils in the west are still far less productive than the fertile soils in the east.

Observing the overall trend in parish production, current national yield is more evenly distributed than in previous time periods (Figure 6). In 1688 and 1844, the data were skewed towards a large portion of parishes having low yield potential and a few having high yield potential. The slight increase in yield potential from 1688 and 1844 could be attributed to the large land reform between the two time periods. During this land reform in the late 18th century, the land that was once owned by the king, church, and manors, and farmed by peasant farmers, was redistributed to the people to allow farmers to cultivate their own land. The increase in agriculture productivity has been associated with the peasant emancipation and the reforms [61]. As the years have progressed, Denmark has become more productive. Current yield is evenly spread across parishes from low to high, with most of the parishes having average yields.

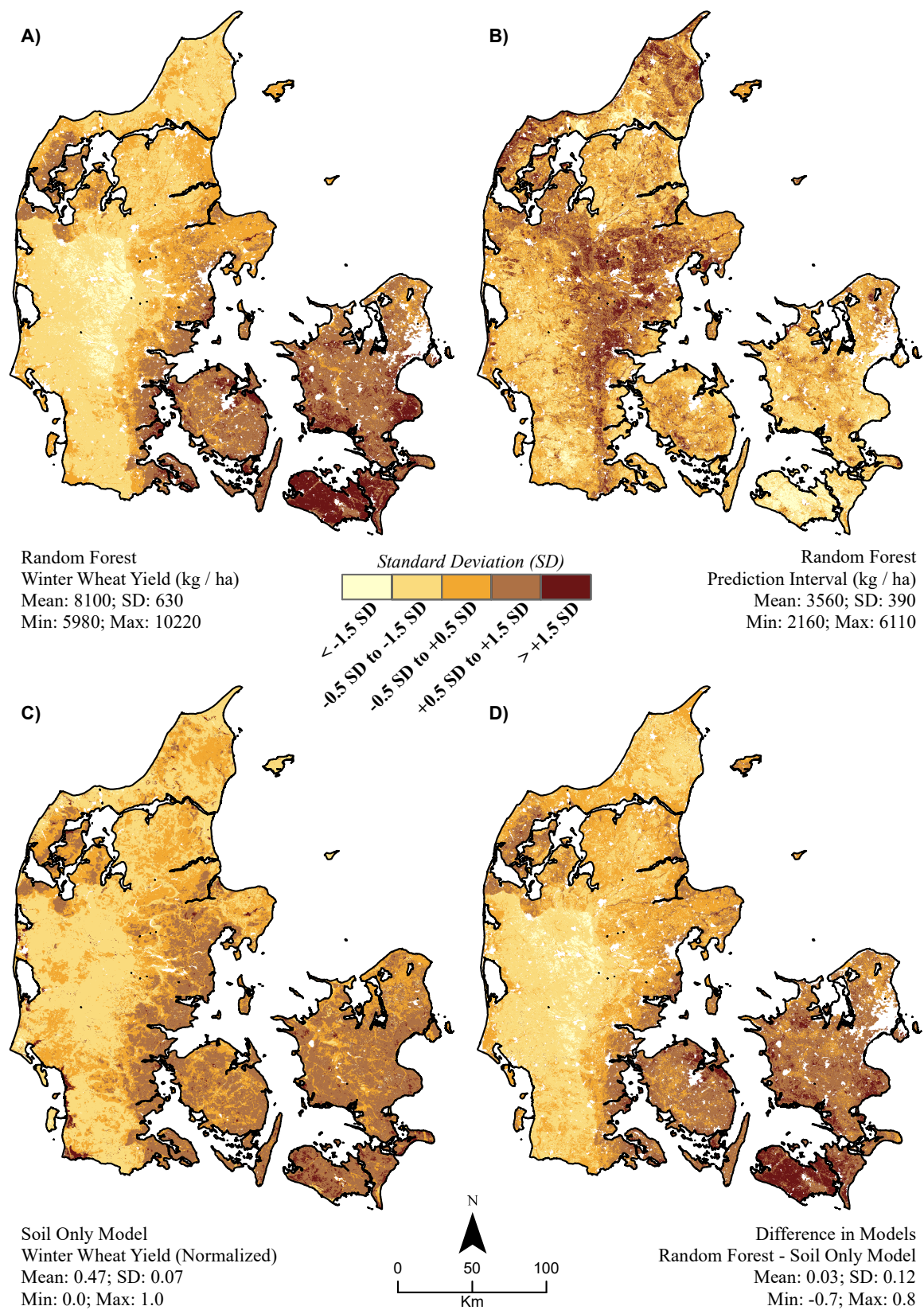


Figure 4. (A) Yield map of winter wheat produced from random forest model. (B) Prediction interval (quantiles 0.9–0.1) of random forest model. (C) Yield map from soil-only model. (D) Comparison between the two models (random forest–soil-only model).

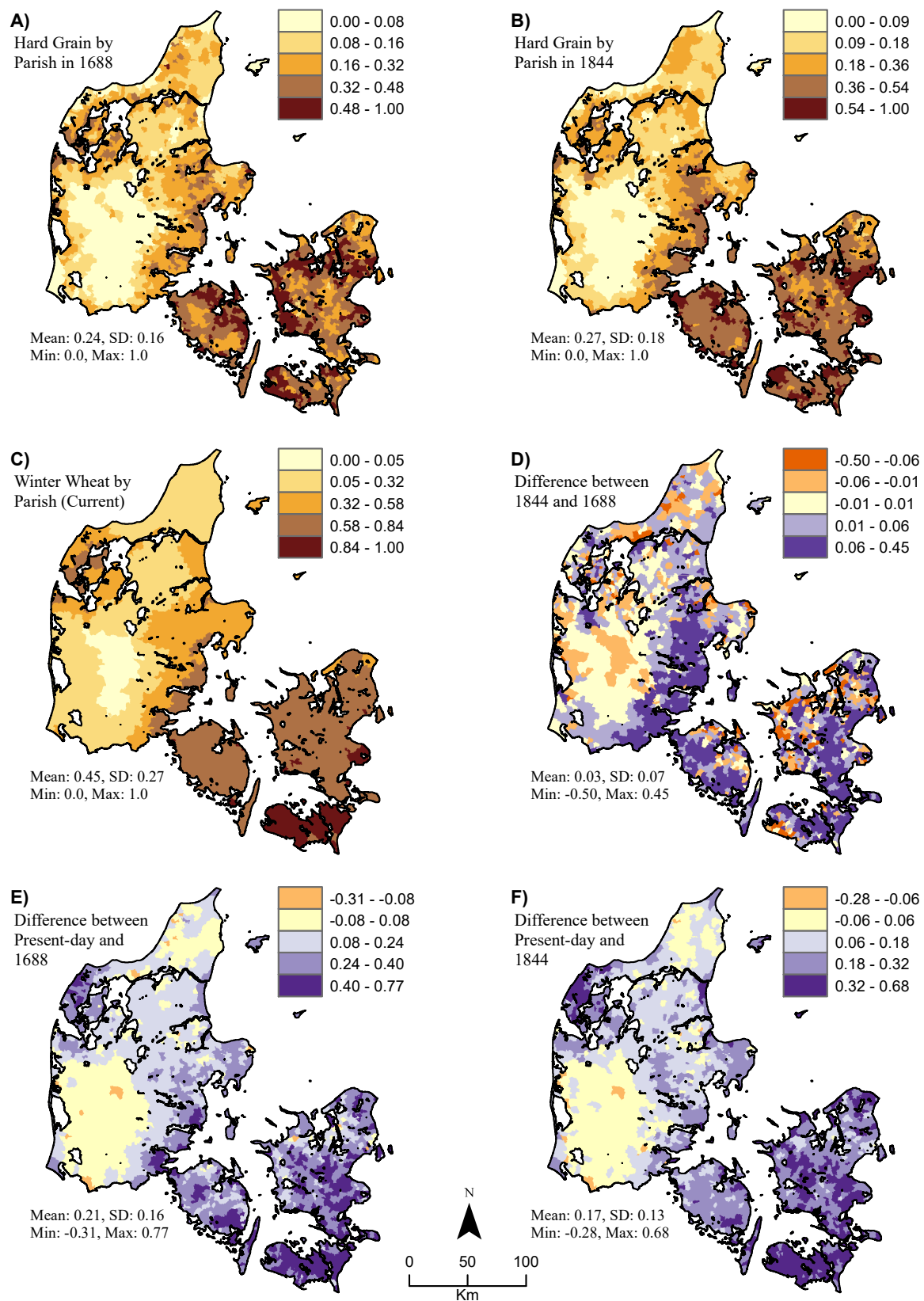


Figure 5. The three time periods compared by parish. (A–C) Normalized values for hard grain in 1688, 1844, and current winter wheat yield, respectively. (D–F) Differences between each time period.

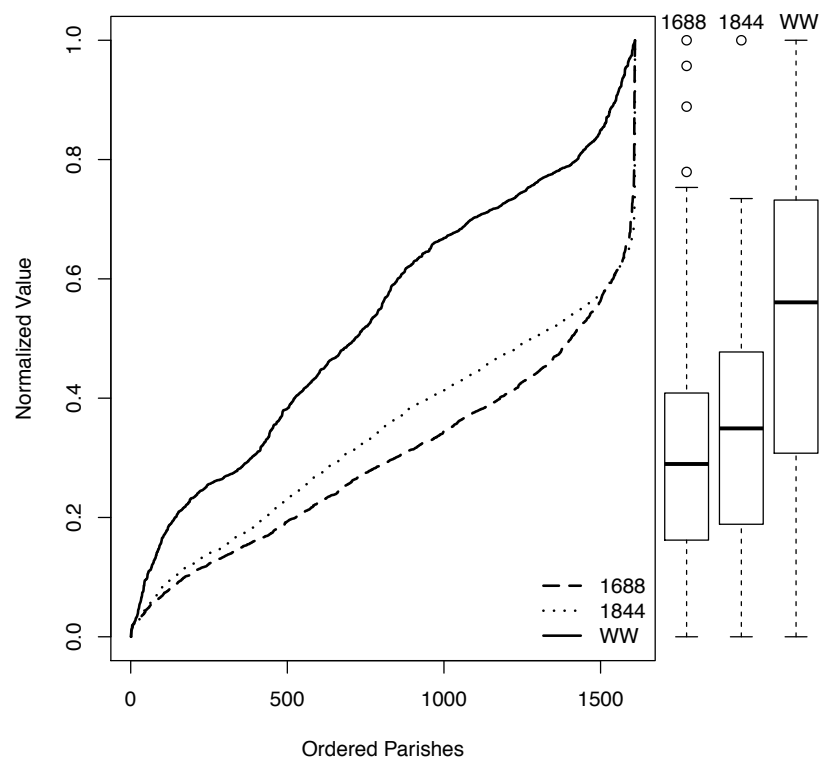


Figure 6. Comparison of the three time periods (WW = current winter wheat yield). Parishes are ordered by normalized values from smallest to largest for each time period separately.

The differences in historical yield and current yield could be attributed to new cultivars, technology spreading, and climate differences. Selective breeding has increased yield through the years with no yield plateau being discovered [62]. The introduction of machinery for ploughing, irrigation, and harvesting, doubled the agricultural capital in Denmark [63]. With drainage pipes and marling being introduced in the mid-19th century, land improved for agriculture, and production quantities grew by 75% [63]. Currently, about 50% of agricultural land is drained in Denmark [64], resulting in more area to be harvested and higher yields compared to historical yield potential. From 1930 to 1970, tile drains were mainly being placed in wetland areas, increasing the amount of arable land in parishes [65,66]. One qualifier to be the lowest class of soil fertility in the historical registers was waterlogged soil [23]. Draining waterlogged soils would cause these soils that once were given a low hard grain value to be given a much higher hard grain value. The loamy soils in southeastern Denmark are now highly productive areas due to the introduction of tile drains. The increase in CO₂ levels through the years can contribute to the increase in yield through time. The CO₂ level during the 18th century was around 280 ppm and current day is around 400 ppm [67]. The change in temperature is also a cause for differences between historical and current production. The temperature during 1875 to 1900 was around 2 °C lower compared to average annual temperature data in Denmark from 2010 to 2017 [68]. Advances in crop genetics, the introduction of new technologies and management practices after the 1844 register, and higher ambient CO₂ levels and temperatures, are major reasons why the current yield values are larger than historical values.

4. Conclusions

The winter wheat yield map of Denmark generated in this study incorporates soil, climate, and topography to produce a more accurate yield prediction than previous efforts. Using a machine learning algorithm and adding climate data were major aspects that helped increase the variation explained in yield predictions compared to the previous statistical model. The previous effort, only using three topsoil properties to model winter wheat yield, was too simple to correctly predict

yield. The random forest model did not predict extreme values well, but the values were centered around the mean. Even though extreme values will not be predicted, the new winter wheat yield map could be applicable for future yield predictions due to the centered values being more reliable predictions. The ease with which machine learning allows additional data also means the map can be updated more readily when new data are collected.

Comparing the current yield to historical yield potentials from the 17th and 19th centuries provided a way to determine the spatiotemporal trends in Danish agriculture. The differences in yield potential between the historical data are minor when comparing the current productivity to historical periods. The advances in agriculture have proven to out-perform what was once thought of as a field's yield potential. The agriculture potential will keep rising as technological breakthroughs continue to occur. Since predicting how technology and new cultivars will change yield potential is challenging, future work should use climate projections when creating yield maps to make the predictions viable in future years. The current yield map, while better than the previous soil-only model, will likely be obsolete in future years because of changes in climate. Thus, the trend in yield changes between time periods should be evaluated so that future yield predictions incorporate temporal trends, generating predictions that last longer into the future.

Author Contributions: Conceptualization: Y.E.R. and M.H.G.; data curation: Y.E.R. (current data) and P.G.M. (historical data); visualization: Y.E.R. and M.B.G.; supervision and funding: M.H.G.; writing—original draft preparation: Y.E.R. and A.B.; writing—review and editing: A.B., P.G.M., M.B.G. and M.H.G.; methodology, software, and validation: Y.E.R. All authors have read and agreed to the published version of the manuscript.

Funding: The study was supported by the ProvenanceDK Project with funding from the Innovation Fund Denmark (grant number 6150-00035B).

Acknowledgments: We are thankful to Leif Knudsen from SEGES and Johannes Jensen from Aarhus University for access to and help with acquiring the data from the Danish National Field Trials. We appreciate the working space provided by Southern Denmark University.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study (in the collection, analyses, or interpretation of data), in the writing of the manuscript, or in the decision to publish the results.

References

1. Ritchie, J.; Godwin, D.; Otter, S. *CERES-Wheat: A User Oriented Wheat Yield Model. Preliminary Documentation*; AGRISTARS Publications: East Lansing, MI, USA, 1984; p. 252.
2. Van Diepen, C.; Wolf, J.; Van Keulen, H.; Rappoldt, C. WOFOST: A simulation model of crop production. *Soil Use Manag.* **1989**, *5*, 16–24. [[CrossRef](#)]
3. Jamieson, P.; Semenov, M.; Brooking, I.; Francis, G. Sirius: A mechanistic model of wheat response to environmental variation. *Eur. J. Agron.* **1998**, *8*, 161–179. [[CrossRef](#)]
4. Wheat Initiative. *Annual Report 2018*; Federal Research Centre for Cultivated Plants: Berlin, Germany, 2018.
5. USDA. *Europe Crop Production Maps for Wheat*; Foreign Agriculture Service at United States Department of Agriculture: Washington, DC, USA, 2015.
6. Satorre, E.; Slafer, G. (Eds.) *Wheat: Ecology and Physiology of Yield Determination*; Haworth Press Technology and Industrial: Binghamton, NY, USA, 1999.
7. Reynolds, M.; Foulkes, M.; Slafer, G.; Berry, P.; Parry, M.; Snape, J.; Angus, W. Raising yield potential in wheat. *J. Exp. Bot.* **2009**, *60*, 1899–1918. [[CrossRef](#)] [[PubMed](#)]
8. Reynolds, M.; Bonnett, D.; Chapman, S.; Furbank, R.; Manés, Y.; Mather, D.; Parry, M. Raising yield potential of wheat. I. Overview of a consortium approach and breeding strategies. *J. Exp. Bot.* **2011**, *62*, 439–452. [[CrossRef](#)] [[PubMed](#)]
9. Parry, M.; Reynolds, M.; Salvucci, M.; Raines, C.; Andralojc, P.; Zhu, X.; Price, G.; Condon, A.; Furbank, R. Raising yield potential of wheat. II. Increasing photosynthetic capacity and efficiency. *J. Exp. Bot.* **2011**, *62*, 453–467. [[CrossRef](#)]
10. Foulkes, M.; Slafer, G.; Davies, W.; Berry, P.; Sylvester-Bradley, R.; Martre, P.; Calderini, D.; Griffiths, S.; Reynolds, M. Raising yield potential of wheat. III. Optimizing partitioning to grain while maintaining lodging resistance. *J. Exp. Bot.* **2011**, *62*, 469–486. [[CrossRef](#)]

11. Miller, M.; Singer, M.; Nielsen, D. Spatial variability of wheat yield and soil properties on complex hills. *Soil Sci. Soc. Am. J.* **1988**, *52*, 1131–1141. [\[CrossRef\]](#)
12. Lobell, D.; Burke, M. Why are agricultural impacts of climate change so uncertain? The importance of temperature relative to precipitation. *Environ. Res. Lett.* **2008**, *3*, 1–8. [\[CrossRef\]](#)
13. Ferrara, R.; Trevisiol, P.; Acutis, M.; Rana, G.; Richter, G.; Baggaley, N. Topographic impacts on wheat yields under climate change: two contrasted case studies in Europe. *Theor. Appl. Climatol.* **2010**, *99*, 53–65. [\[CrossRef\]](#)
14. Kumhalova, J.; Kumhala, F.; Kroulik, M.; Matejkova, S. The impact of topography on soil properties and yield and the effects of weather conditions. *Precis. Agric.* **2011**, *12*, 813–830. [\[CrossRef\]](#)
15. Song, X.; Mitnitski, A.; Cox, J.; Rockwood, K. Comparison of machine learning techniques with classical statistical models in predicting health outcomes. *Medinfo* **2004**, 736–740.
16. Quiza, R.; Figueira, L.; Davim, J. Comparing statistical models and artificial neural networks on predicting the tool wear in hard machining D2 AISI steel. *Int. J. Adv. Manuf. Technol.* **2008**, *37*, 641–648. [\[CrossRef\]](#)
17. Pantazi, X.; Moshou, D.; Alexandridis, T.; Whetton, R.; Mouazen, A. Wheat yield prediction using machine learning and advanced sensing techniques. *Comput. Electron. Agric.* **2016**, *121*, 57–65. [\[CrossRef\]](#)
18. Saeed, U.; Dempewolf, J.; Becker-Reshef, I.; Khan, A.; Ahmad, A.; Wajid, S. Forecasting wheat yield from weather data and MODIS NDVI using random forests for Punjab province, Pakistan. *Int. J. Remote. Sens.* **2017**, *38*, 4831–4854. [\[CrossRef\]](#)
19. Yang, H.; Dobermann, A.; Cassman, K.; Walters, D.; Grassini, P. *Hybrid-Maize: A Simulation Model for Corn Growth and Yield*; University of Nebraska: Lincoln, UK, 2016.
20. Evans, L.; Fischer, R. Yield potential: Its definition, measurement, and significance. *Crop. Sci.* **1999**, *39*, 1544–1551. [\[CrossRef\]](#)
21. Zhang, W.; Kokubun, M. Historical changes in grain yield and photosynthetic rate of rice cultivars released in the 20th century in Tohoku Region. *Plant Prod. Sci.* **2004**, *7*, 36–44. [\[CrossRef\]](#)
22. Grassini, P.; Eskridge, K.; Cassman, K. Distinguishing between yield advances and yield plateaus in historical crop production trends. *Nat. Commun.* **2013**, *4*, 2918. [\[CrossRef\]](#)
23. Pedersen, H. *De danske Landbrug fremstillet para Grundlag af Forarbejderne til Christian V's Matrikel 1688*; Reprotryk for LandBohistorisk Selskab: Copenhagen, Denmark, 1928.
24. Madsen, H.; Nørr, A.; Holst, K. *Atlas over Danmark: Atlas of Denmark*; The Royal Danish Geographic Society: Copenhagen, Denmark, 1992; pp. 1–56.
25. Danish Agriculture and Food Council. *Facts and Figures: Denmark—A Food and Farming Country*; Landbrug and Fødevarer: Copenhagen, Denmark, 2016; pp. 1–56.
26. Schjøning, P.; Jensen, J.; Bruun, S.; Jensen, L.; Christensen, B.; Munkholm, L.; Oelofse, M.; Baby, S.; Knudsen, L. The role of soil organic matter for maintaining crops yields: Evidence for a renewed conceptual basis. *Adv. Agron.* **2018**, *150*, 35–79.
27. Wang, P. *Referenceværdier: Døgn-, Måned- og årsværdier For Regioner Og Hele Landet 2001–2010, Danmark for Temperatur, Relativ Luftfugtighed, Vindhastighed, Globalstråling og Nedbør*; Teknisk Rapport; Danish Meteorological Institute: Copenhagen, Denmark, 2013; pp. 1–14.
28. Schou, A. *Atlas of Denmark I: The Landscapes*; The Royal Danish Geographic Society: Copenhagen, Denmark, 1949; pp. 1–32.
29. Jacobsen, N. Soil map of Denmark according to the FAO-UNESCO Legend. *Geogr. Tidsskr. Dan. J. Geogr.* **1984**, *84*, 93–98. [\[CrossRef\]](#)
30. SEGES. *Nordic Field Trial System*; Landbrug and Fødevarer: Copenhagen, Denmark, 2018.
31. Breiman, L. Random Forest. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
32. Meinshausen, N. Quantile Regression Forests. *J. Mach. Learn. Res.* **2006**, *7*, 983–999.
33. Wright, M.; Ziegler, A. ranger: A fast implementation of random forests for high dimensional data in C++ and R. *J. Stat. Softw.* **2017**, *77*, 1–17. [\[CrossRef\]](#)
34. Adhikari, K.; Minasny, B.; Greve, M.B.; Greve, M.H. Constructing a soil class map of Denmark based on the FAO legend using digital techniques. *Geoderma* **2014**, *214*, 101–113. [\[CrossRef\]](#)
35. Adhikari, K.; Hartemink, A.; Minasny, B.; Kheir, R.; Greve, M.B.; Greve, M. Digital Mapping of Soil Organic Carbon Contents and Stocks in Denmark. *PLoS ONE* **2014**, *9*, e105519. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Adhikari, K.; Kheir, R.; Greve, M.B.; Bøcher, P.; Malone, B.; Minasny, B.; McBratney, A.; Greve, M.H. High-Resolution 3-D Mapping of Soil Texture in Denmark. *Soil Sci. Soc. Am. J.* **2013**, *77*, 860–876. [\[CrossRef\]](#)

37. Møller, A.; Iversen, B.; Beucher, A.; Greve, M. Prediction of soil drainage classes in Denmark by means of decision tree classification. *Geoderma* **2019**, *352*, 314–329. [[CrossRef](#)]
38. Henriksen, H.; Højberg, A.; Olsen, M.; Seaby, L.; Van Der Keur, P.; Stisen, S.; Troldborg, L.; Sonnenborg, T.; Refsgaard, J. *Klimaeffekter På hydrologi og Grundvand—Klimagrundvandskort*; Aarhus University: Aarhus, Denmark, 2012.
39. Adhikari, K.; Kheir, R.; Greve, M.B.; Greve, M.H.; Malone, M.; Minasny, B.; McBratney, A. Mapping soil pH and bulk density at multiple soil depths in Denmark. In *GlobalSoilMap: Basis of the Global Spatial Soil Information System*; Taylor & Francis: London, UK, 2014; pp. 155–160.
40. Jacobsen, P.; Hermansen, B.; Tougaard, L. *Danmarks Digitale Jordartskort 1:25000 Version 4.0*; Geological Survey of Denmark and Greenland (GEUS): Copenhagen, Denmark, 2015.
41. Climate, A. *AGRO Climate: Klimadatabase*; Aarhus University: Aarhus, Denmark, 2018.
42. Scharling, M. *Klimagrid-Danmark Normaler 1961-90 måneds- og årsværdier nedbør 10*10, 20*20 & 40*40 km Temperatur Og Potentiel Fordampning 20*20 & 40*40 km*; Teknisk Rapport; Danish Meteorological Institute: Copenhagen, Denmark, 2000; pp. 1–17.
43. Bergsøe, A. *Den Danske Stats Statistik, Vol. 1-4*; Statistics Denmark: Copenhagen, Denmark, 1853.
44. Møller, P.; Porsmose, E.; Schou, A. *Kulturhistorisk Inddeling Af Landskabet*; Miljø- og Energiministeriet: Copenhagen, Denmark, 1997.
45. Thorup-Kristensen, K.; Cortasa, M.; Loges, R. Winter wheat roots grow twice as deep as spring wheat roots, is this important for N uptake and N leaching losses? *Plant Soil* **2009**, *322*, 101–114. [[CrossRef](#)]
46. Kravchenko, A.; Bullock, D. Correlation of corn and soybean grain yield with topography and soil properties. *Agron. J.* **2000**, *92*, 75–83. [[CrossRef](#)]
47. Riha, S.; Wilks, D.; Simoens, P. Impact of temperature and precipitation variability on crop model predictions. *Clim. Chang.* **1996**, *32*, 293–311. [[CrossRef](#)]
48. Iizumi, T.; Shin, Y.; Kim, W.; Kim, M.; Choi, J. Global crop yield forecasting using seasonal climate information from a multi-model ensemble. *Clim. Serv.* **2018**, *11*, 13–23. [[CrossRef](#)]
49. Kristensen, K.; Schelde, K.; Olesen, J. Winter wheat yield response to climate variability in Denmark. *J. Agric. Sci.* **2011**, *149*, 33–47. [[CrossRef](#)]
50. Webber, H.; Ewert, F.; Olesen, J.; Muller, C.; Fronzek, S.; Ruane, A.; Bourgault, M.; Martre, P.; Ababaei, B.; Bindi, M.; et al. Diverging importance of drought stress for maize and winter wheat in Europe. *Nat. Commun.* **2018**, *9*, 1–10. [[CrossRef](#)] [[PubMed](#)]
51. Ørum, J.; Jensen, J.; Thomsen, I.; Knudsen, L. *Optimal Kvælstoftildeling til Korn: Responsfunktioner og Metode til Produktionsøkonomisk Analyse af Landsforsøgene Med Kvælstoftildeling til Vinterhvede og vårbyg Samt Kvalitetssikring af Normudvalgets Indstillinger*; IFRO Rapport, Nr. 281; Institut for Fødevarer- og Ressourceøkonomi, Københavns University: Copenhagen, Denmark, 2019.
52. Kern, A.; Barcza, Z.; Marjanovic, H.; Árendas, T.; Fodor, N.; Bonis, P.; Bogнар, P.; Lichtenberger, J. Statistical modelling of crop yield in Central Europe using climate data and remote sensing vegetation indices. *Agric. For. Meteorol.* **2018**, *260*, 300–320. [[CrossRef](#)]
53. Verrillo, F.; Badeck, F.; Terzi, V.; Rizza, F.; Bernardo, L.; Di Maro, A.; Fares, C.; Zaldei, A.; Miglietta, F.; Moschella, A.; Bracale, M.; Vannini, C. Elevated field atmospheric CO₂ concentrations affect the characteristics of winter wheat (cv. Bologna) grains. *Crop. Pasture Sci.* **2017**, *68*, 713–725. [[CrossRef](#)]
54. Broberg, M.; Høgy, P.; Feng, Z.; Pleijel, H. Effects of elevated CO₂ on wheat yield: Non-linear response and relation to site productivity. *Agronomy* **2019**, *9*, 243. [[CrossRef](#)]
55. Lin, L. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* **1989**, *45*, 255–268. [[CrossRef](#)]
56. Fleiss, J. *Statistical Methods for Rates and Proportions*, 1st ed.; John Wiley & Sons: Hoboken, NJ, USA, 1981.
57. Green, T.; Salas, J.; Martinez, A.; Erskine, R. Relating crop yield to topographic attributes using Spatial Analysis Neural Networks and regression. *Geoderma* **2007**, *139*, 23–37. [[CrossRef](#)]
58. Jeong, J.; Resop, J.; Mueller, N.; Fleisher, D.; Yun, K.; Butler, E.; Timlin, D.; Shim, K.; Gerber, J.; Reddy, V.; Kim, S. Random forests for global and regional crop yield predictions. *PLoS ONE* **2016**, *11*, e0156571. [[CrossRef](#)]
59. Barbero, F.; Lopez, G.; Batlles, F. Determination of daily solar ultraviolet radiation using statistical models and artificial neural networks. *Ann. Geophys.* **2006**, *24*, 2105–2114. [[CrossRef](#)]
60. Rothe, C. Beretning om den i Aaret 1844, for Kongeriget indførte nye Skyldsætnings Væsen og Historie. Copenhagen. 1844.

61. Barton, H. The Danish agrarian reforms, 1784–1844, and the historians. *Scand. Econ. Hist. Rev.* **1988**, *36*, 46–61. [[CrossRef](#)]
62. Cox, T.; Shroyer, J.; Ben-Hui, L.; Sears, R.; Martin, T. Genetic improvement of agronomic traits of hard red winter wheat cultivars from 1919 to 1987. *Crop. Sci.* **1988**, *28*, 756–760. [[CrossRef](#)]
63. Christensen, J. *Rural Denmark 1750–1980*; The Central Co-Operative Committee of Denmark: Copenhagen, Denmark, 1983.
64. Olesen, S. *Kortlægning af Potentielt Dræningsbehov På Landbrugsarealer Opdelt efter Landskabslement, Geologi, Jordklasse, Geologisk Region Samt Høj/lavbund*; Det Jordbrugsvidenskabelige Fakultet Markburg Nr. 1; Aarhus University: Aarhus, Denmark, 2009.
65. Hansen, B.; Olesen, S.; Ernsten, V. Naturlig Dræning og Grundvandsdannelse. In *Afprøvning af Undersøgelsesmetoder med Henblik på Etablering af et Zoneringskoncept for Danske Lejorde: Statusrapport*; Geological Survey of Denmark and Greenland, Appendix 10; Geological Survey of Denmark and Greenland (GEUS): Copenhagen, Denmark, 2004.
66. Madsen, H. Drænrørets indførelse og betydning i et landbrugs- og miljømæssigt perspektiv. In *Det Fremmede Som Historisk Drivkraft: Danmark Efter 1742*; Det Kongelige Danske Videnskabernes Selskab: Copenhagen, Denmark, 2010; pp. 158–165.
67. Lindsey, R. *Climate Change: Atmospheric Carbon Dioxide*; National Oceanic and Atmospheric Administration: Copenhagen, Denmark, 2020.
68. Cappelen, J. *Denmark-DMI Historical Climate Data Collection 1768–2017*; DMI Report 18-02; Danish Meteorological Institute: Copenhagen, Denmark, 2018.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).