agronomy

MDPI

Article

Soil Dynamics and Crop Yield Modeling Using the MONICA Crop Simulation Model and Time Series Forecasting Methods

Islombek Mirpulatov [1], Mikhail Gasanov [1,*] and Sergey Matveev [2,3]

[1] Applied Artificial Intelligence Center, Skolkovo Institute of Science and Technology, Moscow 121205, Russia; i.mirpulatov@skoltech.ru
[2] Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Moscow 119991, Russia; matseralex@cs.msu.ru
[3] Marchuk Institute of Numerical Mathematics, Russian Academy of Sciences (RAS), Moscow 119333, Russia
* Correspondence: mikhail.gasanov@skoltech.ru

Abstract: Crop simulation models are an important tool for assessing agroecosystem performance and the impact of agrotechnologies on soil cover condition. However, the high uncertainty and labor intensiveness of long-term weather forecasting limits the applicability of such models. A possible solution may be to use time series forecasting models (SARIMAX and Prophet) and artificial neural-network-based technologies (Neural Prophet). This work compares the applicability of these methods for modeling soil condition dynamics and agroecosystem performance using the MONICA simulation model for Voronic Chernozems in the Kursk region of Russia. The goal is to determine which weather indicators are most important for the yield forecast and to choose the most appropriate methods for forecasting weather scenarios for agricultural modeling. Crop rotation of soybean and sugar beet was simulated, with agricultural techniques and fertilizer usage considered as factors. We demonstrated the high sensitivity of aboveground biomass production and soil moisture dynamics to daily temperature fluctuations and precipitation during the vegetation period. The dynamics of the leaf area index and nitrate content showed less sensitivity to the daily fluctuations of temperature and precipitation. Among the proposed forecasting methods, both SARIMAX and the Neural Prophet algorithm demonstrated the ability to forecast weather to model the dynamics of crop and soil conditions with the highest degree of approximation to actual observations. For the dynamic of the crop yield of soybean, the SARIMAX model exhibited the most favorable coefficient of determination, R^2, while for sugar beet, the Neural Prophet model achieved superior R^2 levels of 0.99 and 0.98, respectively.

Keywords: crop simulation model; autoregressive models; soil dynamics modeling

Citation: Mirpulatov, I.; Gasanov, M.; Matveev S. Soil Dynamics and Crop Yield Modeling Using the MONICA Crop Simulation Model and Time Series Forecasting Methods. *Agronomy* **2023**, *13*, 2185. https://doi.org/10.3390/agronomy13082185

Academic Editors: Wen-Shin Lin and Yun Yang Chao

Received: 21 July 2023
Revised: 13 August 2023
Accepted: 17 August 2023
Published: 21 August 2023

1. Introduction

Global population growth is driving the intensification of agricultural production and soil use. The increase in the consumption of agricultural products is determined by many factors, such as the availability of resources, water, and fertilizers, and environmental pollution. Global climate change brings even more uncertainty to this process [1,2]. The efficient use of resources and forecasting the impact of various stress factors on agricultural production are important to ensure food security and reduce the negative impact on the environment [3]. One of the approaches to planning economic activity and forecasting the impact on agroecosystems is the use of dynamic simulation models of agroecosystem performance [4–6].

Dynamic crop simulation models are used to predict crop yields and estimate potential yields in different regions [7] and are also implemented in the decision support systems for farmers [8]. Modeling requires the data on soil cover, environmental conditions, weather observations, and crop management practices including sowing and harvesting dates,

and the types and amounts of fertilizers used. The most commonly used crop simulation models such as DSSAT [9], APSIM [10], STICS [11], AgroTool [12], and MONICA [13] include soil, climate, and crop properties modules.

The description of soil processes included in most simulation models allows them to be used to predict soil conditions and the dynamics of soil components [14]. In agroecosystem modeling tasks, it is of interest to predict the dynamics of nutrient elements and evaluate their loss as a result of washing out from the soil profile [15,16]. Most models analyze soil profiles by separate horizons, between which substance and energy exchange takes place [17]. The models mostly take into account heat exchange and moisture transfer, as well as C and N cycles [14].

Daily weather observations are necessary to obtain reliable yield forecasts. Many factors limit the possibility of obtaining a long-term weather forecast with daily resolution; for this reason, various methods of weather scenario generation are used in the models. Using a forecast of global climate models as input for crop process-based models is problematic for two reasons: the low spatial resolutions of such models and the notable fluctuations and deviations in crucial processes for crop modeling at the daily time step, such as precipitation, droughts, and others [18,19]. Some of the most commonly used approaches to tackle this limitations are stochastic weather generators [20] and statistical time series forecasting methods [21].

The recent progress in the development of artificial intelligence technologies has made it possible to apply deep neural networks for time series prediction, including weather data [22,23]. The convolutional neural network was used to downscale the climate model scenario from the CMIP6 [24] project, and it can be helpful for agricultural modeling tasks [25]. Machine learning methods are also used for seasonal forecasting, and it is worth noting that the authors have attempted to use interpreted modeling approaches [26,27]. The hybrid use of data-driven and physical-based weather and climate forecast models has become popular in the last decade [28,29].

While stochastic weather generators are quite often used in agroecosystem modeling, machine learning and neural networks have hardly ever been used for weather forecasting in the simulation models. In this paper, we analyze the possibility of using the crop simulation models together with the advanced time series forecasting methods. We also examined how sensitive crop yield models are to the quality of weather forecasts and estimated which weather components impact crop yield more. We assumed that the quality of the weather forecast affects the quality of the estimates of the crop dynamics in different ways. A location in the Kursk Oblast, Russia, was chosen to perform numerical modeling. Modeling was carried out using the MONICA simulation model for sugar beet (sugar beet, *Beta vulgaris*) and soybean (soyabean, *Glycine max*) in 2016 and 2017.

The purpose of this work was to investigate whether it is possible to use time series forecasting methods for the modeling of soil condition dynamics and agroecosystem performance using the MONICA simulation model.

The main contributions of this paper are:

- We demonstrated the performance of time series models to address weather scenario generation for crop modeling;
- We estimated the efficiency of classical and neural network time series methods for crop and soil modeling;
- We implemented a crop yield forecast based on time series models and a crop simulation model.

2. Materials and Methods

2.1. Research Object

An agricultural field located in the Kursk Oblast, Sovetsky District, Kshensky Settlement, was chosen as a pilot research object. The terrain is a rolling plain crossed by ravines with smooth slopes and shapeless depressions. Climatic conditions correspond to the moderately cold climate with the average annual precipitation of 350–570 mm and the

average annual ambient temperature of 5.4 °C. The average temperatures are −7.2 °C in January and +20.1 °C in June. The vegetation period lasts for 188 days. The area of the field plot used in modeling was 48 ha (Figure 1).

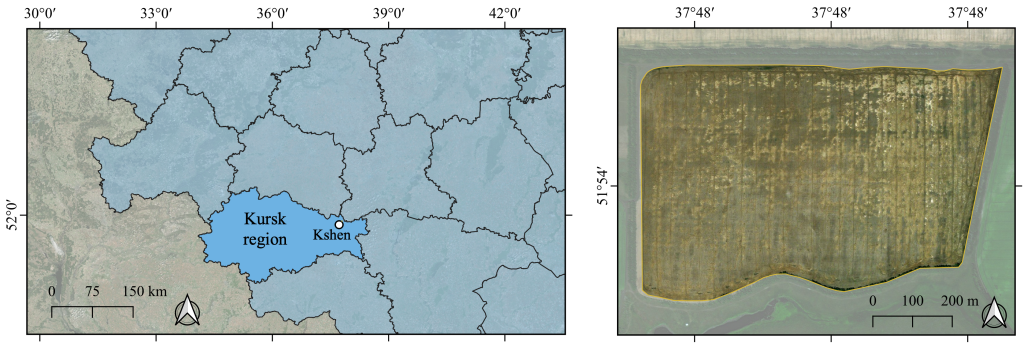

Figure 1. Study area.

Soil was classified as Voronic Chernozem according to the World Reference Base for Soil Resources [30]. An uneven four-field crop rotation was used: barley (2012)—winter wheat (2013)—sugar beet (2014)—winter wheat (2015)—soybean (2016)—sugar beet (2017). Soil samples were collected from the depths of 10, 30, and 80 cm at 51 locations within the field boundaries (N 51 54′8′, E 37 47′34′) in October 2016. The key soil characteristics used in modeling are summarized in Table 1.

Table 1. Soil characteristics used as input data in the modeling.

Horizon Depth, cm	Humus, %	Sand Fraction, %	Clay Fraction, %	pH	C:N	Bulk Density, kg/m^3
0–30	5.1	37	9	6.2	12.4	1126
30–50	3.3	45	8.8	6.4	10.5	1056
50–70	2.3	49	8.6	6.6	8.8	1010
70–200	1.6	5.3	9.4	6.7	7.5	1113

2.2. Weather Data

Prediction of Worldwide Energy Resource (POWER) data, which is a publicly available NASA project, was used as input data for generating weather scenarios and training machine learning models. The POWER system is a common source of meteorological data in agricultural modeling, including crop yield modeling [31,32], crop disease spread modeling [33], field irrigation optimization [34], and other tasks [35,36]. The data are created by the assimilation of satellite observations into a climate model (Goddard Earth Observing System climate model) and cover the entire land surface [37]. NASA POWER data exhibit a high level of concordance between the obtained estimates and observations at individual weather stations. Other advantages of this data source are the high spatial resolution (1° × 1° degrees) and high speed of updating with the lag of only 5–7 days. The observations starting from 1983 are available at daily resolution. The following observations available in POWER were used as input parameters in modeling: minimum temperature (T2M_MIN, °C), maximum temperature (T2M_MAX, °C), average temperature (T2M, °C), precipitation (PRECTOTCORR, mm), incoming solar radiation (ALLSKY_SFC_SW_DWN, W/m^2), wind speed (WS2M, m/s), and relative humidity (RH2M, %). The NASA POWER data were obtained for the Kshensky Settlement, Kursk Oblast (N 51 54′8′, E 37 47′34′), for the years 2010–2017 in CSV format.

2.3. Time Series Forecasting Methods

Here we describe the time series methods used to forecast weather scenarios. Conventional methods and methods based on neural networks were used.

As a baseline we used grouping by mean approach. Grouping by mean value is the simplest method of forecasting a time series with consists in averaging weather observations over several years. The data were grouped by date and averaged over 10 years. As a result, the averaged data were fed to the input of the simulation model.

SARIMAX (Seasonal Autoregressive Integrated Moving Average with exogenous regressors) is among the most common methods for analyzing and forecasting time series [38,39]. SARIMAX combines autoregression (AR), the moving average (MA), and integration (I) for time series modeling. The AR model is used to describe the dependence of the current value on past values, the MA model is used to describe the dependence of the current value on the random errors of the past values, and the I model is used to eliminate non-stationarity of the series:

$$AR(p): \quad y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \tag{1}$$

where y_t—stationary time series with zero mean, $\phi_1, \ldots, \phi_p$—constants $(\phi_p \neq 0)$, ε_t—Gaussian white noise with zero mean and constant variance σ_ε^2.

$$MA(q): \quad y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}, \tag{2}$$

where y_t—stationary time series with zero mean, $\theta_1, \ldots, \theta_q$—constants $(\theta_q \neq 0)$, and ε_t—Gaussian white noise with zero mean and constant variance σ_ε^2.

In addition, SARIMAX includes the seasonality component, which allows us to take into account the repeating cycles in the data. For this purpose, the seasonality parameters are used, which define the periodicity. The key idea of SARIMAX lies in adding exogenous factors to the ARIMA model. To do this, additional regressors are introduced into the model that can affect the target time series. The general formula of SARIMAX is as follows:

$$Y_t = \beta_0 + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \sum_{i=1}^{P} \Phi_i Y_{t-i}^{(s)} + \sum_{i=1}^{Q} \Theta_i \epsilon_{t-i}^{(s)} + X_t^T \beta + \epsilon_t \tag{3}$$

where Y_t—value at the time point t, β_0—constant, p—order of autoregression, ϕ_i—autoregression coefficient, q—order of moving average, θ_i—moving average coefficient, P—order of seasonal autoregression, Φ_i—seasonal autoregression coefficient, Q—order of seasonal moving average, Θ_i—seasonal moving average coefficient, $Y_{t-i}^{(s)}$—value at the $t-i$ time point in season s, $\epsilon_{t-i}^{(s)}$—error at the time point $t-i$ in season s, X_t—vector of exogenous regressors at the time point t, β—vector of regression coefficients, and ϵ_t—error at the time point t.

In this way, SARIMAX takes into account seasonality, autocorrelation, as well as the effects of exogenous variables on the target time series. The maximum likelihood method is used to determine the optimal model parameters.

Prophet is a time series forecasting procedure [40]. It is based on the approach that allows the modeling of various time series with a non-linear trend component, seasonality, and holiday effects.

The model can automatically detect changes in trend and seasonality, as well as take into account additional regressors. In addition, the library can provide model interpretation, which makes it possible to analyze the contribution of each component to the forecast. Prophet is built on the basis of the additive time series model:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon(t) \tag{4}$$

where $g(t)$—trend function, $s(t)$—seasonality function, $h(t)$—holiday function, or anomalous data, and $\varepsilon(t)$—error function.

In Prophet, an individual model is fitted for each component:

$$g(t) = X(t)\beta + \gamma(t), s(t) = \sum_{i=1}^{N} a_i \cos\left(\frac{2\pi i t}{P}\right) + b_i \sin\left(\frac{2\pi i t}{P}\right), h(t) = \sum_{j=1}^{M} k_j I\left(t \in \left[s_j, e_j\right]\right), \tag{5}$$

where $X(t)$—regressor matrix, β—parameter vector, γ—smoothed seasonality function, N—number of seasonality components, P—season length, a_i, b_i—seasonality coefficients, M—number of holidays, k_j—holiday effect strength, and s_j, e_j—start and end of the holiday period, respectively.

Neural Prophet is a time series forecasting method based on neural networks [41]. It is an extension of the Prophet method and employs the same idea of time series decomposition into trend, seasonality, and holidays. However, unlike Prophet, Neural Prophet does not rely on the linear models to model trend and seasonality but utilizes neural networks instead.

The model is based on the LSTM (long short-term memory) architecture [42], which allows modeling dependencies in the data based on past values. In the Neural Prophet model, trend is modeled using neural networks, which can manage non-linear dependencies between the time points. Seasonality and anomalies are also modeled based on neural networks, which allows us to capture more complex patterns in the data. The model can also include exogenous variables to account for additional factors affecting the time series.

2.4. Forecast Accuracy Metrics

There are a number of quality indices that are mostly used to evaluate the quality of time series prediction achieved by different models [43]. The following quality indices were used in this work: Mean Squared Error (MSE), Relative Root Mean Squared Error (RRMSE), and R-squared (R^2). The quality indices were calculated as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{6}$$

$$\text{Relative RMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2}}, \tag{7}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \tag{8}$$

where y_i—actual value of the target variable in the observation i, $\hat{y}_i$—predicted value of the target variable in the observation i, $\bar{y}$—mean value of the target variable, and n—total number of observations.

Dynamic time warping (DTW) was also used to estimate forecasting accuracy. DTW measures the distance between two time series, accounting for possible distortion shifts and different data scales [44]. This metric provides a more accurate measure of similarity between the time series than the simple Euclidean distance. The approach is based on finding the optimal correspondence between the points in the two time series so that the resulting distance between them is minimal. DTW between the two time series is estimated as follows:

$$DTW(X,Y) = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m}(X_i - Y_j)^2}, \tag{9}$$

where X and Y—two time series with the lengths of n and m, respectively.

The DTW (X, Y) procedure computes the Euclidean distance between each point in the time series X and Y and at each time step, and then finds the optimal correspondence between the points using dynamic programming.

2.5. Crop Model MONICA

MONICA [13] was chosen as the agroecosystem model. MONICA is a one-dimensional dynamic process-oriented simulation model of plant production, which describes the transport and biogeochemical turn-over of carbon, nitrogen, and water in agroecosystems.

Using a dynamic iterative algorithm with a daily time step, MONICA simulates the most important processes in the soil and plant, taking into account direct and feedback relations. MONICA is the successor of the previously developed HERMES and AGROSIM models [45,46]. Schematic representation of the MONICA model is presented in Figure 2.

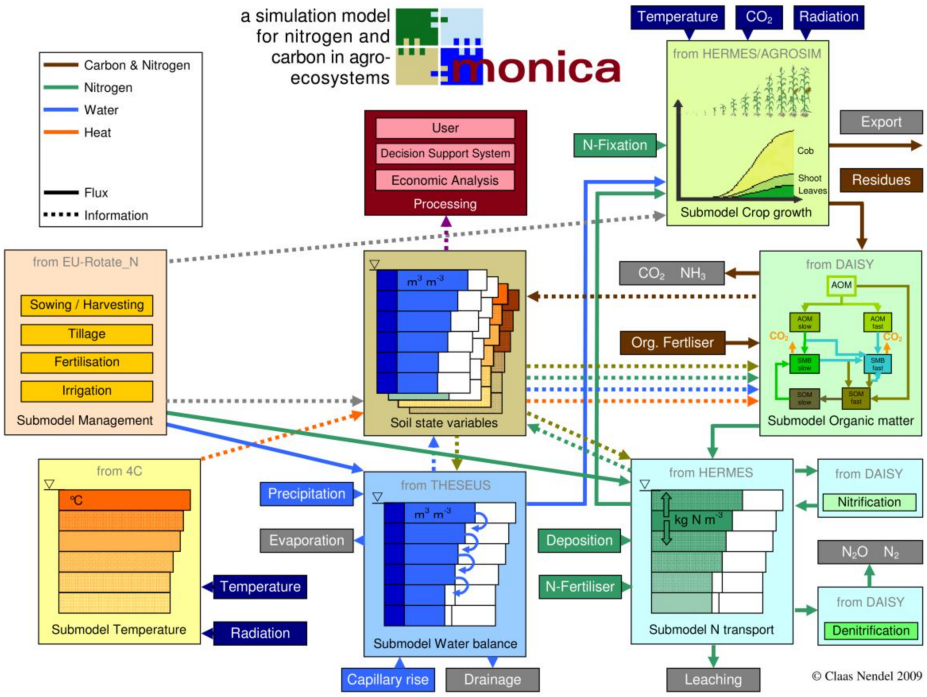

Figure 2. Schematic representation of the MONICA model. The model includes several submodels for describing various processes in soil and culture. Source—documentation of the software implementation of the model [13].

MONICA significantly extended the functionality of its predecessors by adding a full carbon cycle in the agroecosystem, which is a prerequisite for the study of carbon dynamics under climate change. Towards this end, a simple algorithm implemented in HERMES for estimating nitrogen mineralization from soil organic matter was replaced by a comprehensive approach used in the Danish DAISY model [47], which also takes into account the dynamics of soil microbial biomass. MONICA takes into account the impact of the atmospheric CO_2 concentration on photosynthesis and transpiration in crops, as well as the heat stress and possible frost kill of winter crops. The algorithms for calculating plant growth and development implemented in MONICA have been validated for both temperate and tropical latitudes. These experiments identified parameter sets for most typical crops produced in these regions taking into account varietal diversity [48,49]. A sensitivity analysis of MONICA was carried out concerning crop parameters [50] and soil properties [51].

2.6. Numerical Experiments

NASA POWER weather data for the period from 2010 to 2017 were used to build the predictive models of weather scenarios. Two types of experiments were performed to assess forecast accuracy: in the first case, the data were split into a training sample including the years 2010–2015 and a forecast was obtained for 2016 (Figure 3). In the second case, the data were split into a training sample including the years 2011–2016, and a prediction was made for 2017.

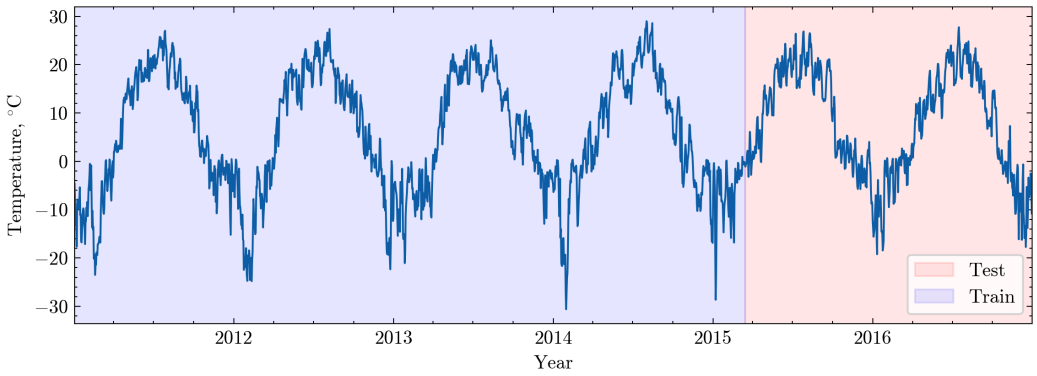

Figure 3. Splitting weather observations for model training into training and test samples.

Then, the SARIMAX, Prophet, and Neural Prophet forecasting models were trained using the available sample, with the last year of weather observations being used for model validation and hyperparameter tuning. The predictions made for 2016 were used to model soybean growth and development, while the results obtained for 2017 were used to model sugar beet growth and development. An outline of the experiments and used software for modeling is illustrated in Figure 4.

To simulate plant growth dynamics and soil condition, the MONICA model was used in this work, for which the input data were weather forecasts and the soil parameters obtained from field surveys (Table 1). Numerical experiments were conducted for two crops: soybean in 2016 and sugar beet in 2017. For both crops, ploughing to a depth of 25–27 cm was performed prior to sowing. The description of the main agrotechnological events for crops is presented in Table 2.

Table 2. Crop rotation and fertilization.

Crop	Tillage (cm)	Fertilization (Date–kgN/ha)	Sowing Date	Harvesting Date
Soyabean	25–27 cm	15.04–50	15 April 2016	9 October 2016
Sugar beet	25–27 cm	23.05–90; 15.06–70	2 May 2017	1 November 2017

The following MONICA model variables were used as target parameters for the studied crops: yield (kg/ha), aboveground biomass (kg/ha), ecosystem carbon exchange (NEE, kgC/ha), and leaf area index (LAI, –). Organic matter carbon dynamics (kgC/ha), nitrate leaching rate (kgN/ha), NO_3 dynamics (kgN/ha), and soil moisture were considered to compare soil condition dynamics.

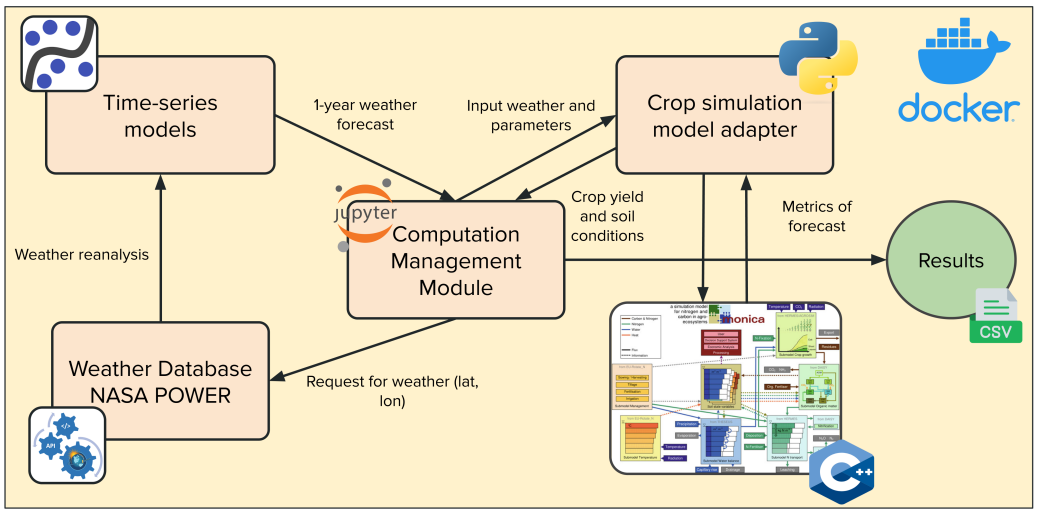

Figure 4. Outline of experiments and used software for modeling.

To assess the quality of the obtained time series of plant growth dynamics and soil condition, the abovementioned quality indices were used, and the forecasts obtained using MONICA based on reanalysis data and experimental data were compared.

3. Results and Discussion

3.1. Forecasting Weather Scenarios

The temperature modeling results are shown in Figure 5. As can be seen on the plot, none of the proposed models can fully describe the weather data. The Prophet and Neural Prophet models smooth the data, resulting in no sharp changes in temperature compared to the real data. If we look at the observations from July to August 2017, it can be seen that the temperature values range from 13 to 26 °C , while the values predicted by the models range from 17 to 20 °C. Most models performed well at capturing the general trend in the data but were rather poor in reproducing the stochasticity of weather events. On the other hand, the grouping by mean and SARIMAX models represented stochasticity; however, the predicted temperature values differed significantly from the real data.

In the case of precipitation modeling, it should be noted that the observations were essentially stochastic and their prediction was nearly impossible. Prophet and Neural Prophet gave smoother results as in the case with temperature but missed the extreme events such as heavy rainfall in June. The grouping by mean model and SARIMAX retained stochasticity, but the peaks predicted by them did not exceed 10 mm/day, while in the real observations, the peaks reached 30 mm/day.

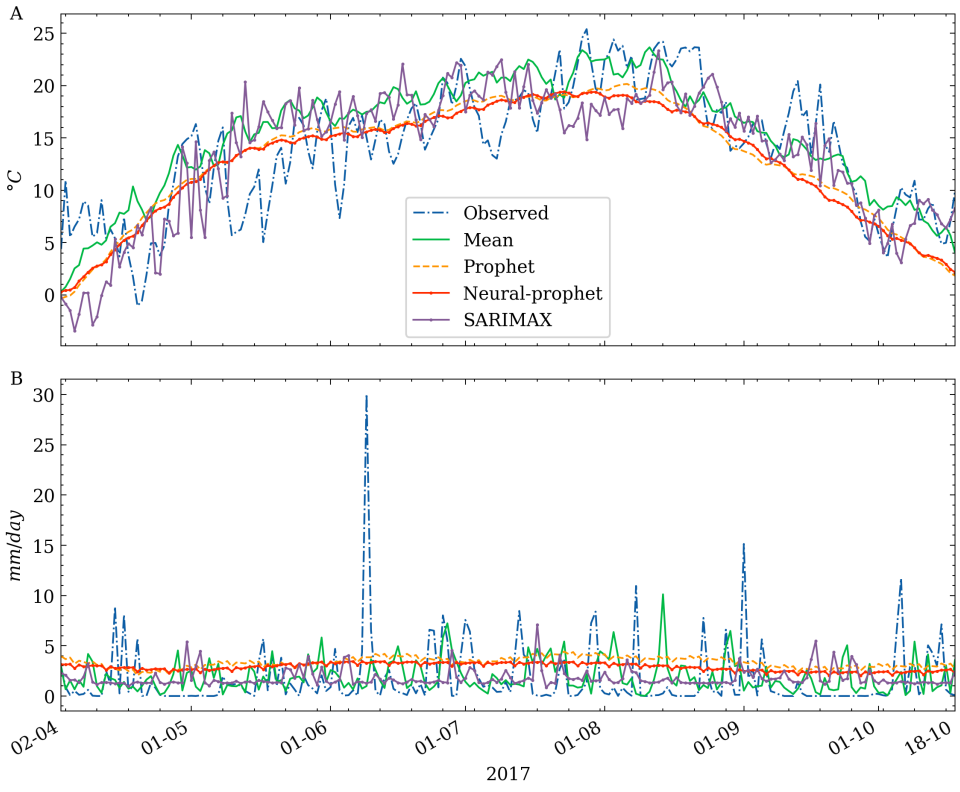

Figure 5. Weather scenario forecasting for average temperature (**A**) and precipitation (**B**) for 2017 using different methods.

The weather conditions forecast accuracy indices are summarized in Table 3. The lowest relative error was achieved by the Prophet class models. However, even in their case, the RRMSE for precipitation and wind speed predictions was up to 99%. The coefficient of determination for precipitation and wind speed forecasts was negative for all models. A negative value can be obtained when the model fits the data worse than a simple constant model, or when the predicted dependent variable values deviate strongly from the actual

values in the sample. The best values for the other variables were also obtained by the Prophet class models. When using the dynamic time warping algorithm to predict average temperature and precipitation, grouping by mean showed the best results, while in the case of solar radiation and wind speed, the best results were achieved with SARIMAX.

Table 3. Weather variables forecast metrics.

	Average Temperature			Precipitation			Solar Radiation			Wind		
	RRMSE	**R^2**	**DTW**	**RRMSE**	**R^2**	**DTW**	**RRMSE**	**R^2**	**DTW**	**RRMSE**	**R^2**	**DTW**
Baseline	0.42	**0.82**	**47.73**	1.05	–0.13	**64.77**	0.52	0.72	50.51	1.00	–0.01	16.67
SARIMAX	0.53	0.71	49.84	1.04	–0.10	68.02	0.57	0.66	**46.72**	1.06	–0.14	**15.88**
Prophet	0.42	0.81	61.20	**0.99**	–0.11	80.34	**0.49**	**0.75**	66.57	0.92	**0.13**	21.94
Neural Prophet	**0.41**	**0.82**	61.50	**0.99**	**–0.05**	80.06	0.50	0.74	66.95	**0.91**	0.10	22.64

Bold indicates best model performance.

Based on the obtained results, it can be concluded that the Prophet and Neural Prophet models are best suited for predicting temperature and solar radiation, as they have the lowest relative error. However, when forecasting precipitation and wind speed, the grouping by mean and SARIMAX models are more suitable as they retain the stochasticity of the data. In general, it is recommended to use Neural Prophet to achieve the best results in time series forecasting, as it demonstrates the highest quality of prediction combined with the good performance for all the variables considered.

3.2. Crop Growth and Development Modeling

The simulations resulted in the growth and development scenarios for soybean and sugar beet for the different forecasts obtained with the machine-learning-based models. The results for each crop were compared considering the dynamic changes in the main crop growth indicators at one-day intervals. Additionally, quality indices were calculated based on the correspondence between the dynamics of the biophysical crop growth indices obtained using different forecasting methods and based on the real weather observations. The growth curves and dynamics of the key biophysical parameters for soybean are provided in Figure 6. The yield dynamic differs considerably between the forecasts obtained with different models. As can be seen in Figure 6A, Neural Prophet and the grouping by mean model both forecast earlier crop maturation than is predicted based on the observed weather conditions. Prophet underpredicts the total yield by about two times. SARIMAX gives the dynamics that are the closest to the real observations for the whole vegetative period. The aboveground biomass dynamics are similar to the yield dynamics. In this case, SARIMAX also produces the dynamics that are the closest to those observed, and also allows values to be obtained that are the closest to the observed estimate of the aboveground biomass at the end of the season.

Ecosystem carbon exchange shows the lowest values in the period from June to August, which is associated with plant biomass growth. Prophet and Neural Prophet give smoother NEE dynamics compared to the actual observations. SARIMAX produces the dynamics that are closer to those observed than all other forecasts, including at the end of the vegetative period when all other models overestimate NEE.

Leaf area index is an important characteristic used to assess crop growth and development during a season. This index is a frequently used variable in the tasks of remote sensing data assimilation into simulation models with the purpose of improving their predictive properties. Therefore, the ability of the model to predict its values during the season is important in assessing its quality. As can be seen on the plot, SARIMAX allows the estimation of LAI dynamics during the season with high accuracy, while the other models have the LAI peak shifted towards the start of the season.

The total soybean yield was 3758 kg/ha in the case of the real weather observations. The closest yield estimate, 3184 kg/ha, was obtained when weather observations were averaged over 10 years, while the worst forecast, 1964 kg/ha, was produced by Prophet. All forecast variants underestimated the final yield. One of the possible reasons for this phenomenon may be the poor quality of precipitation forecasting by the models, which

ultimately leads to the underestimation of precipitation amounts over the season. On the other hand, in the grouping by mean model, the total precipitation was 275 mm, which is almost two times lower than the observed value (504 mm), while the predicted total yield is only lower by 5%.

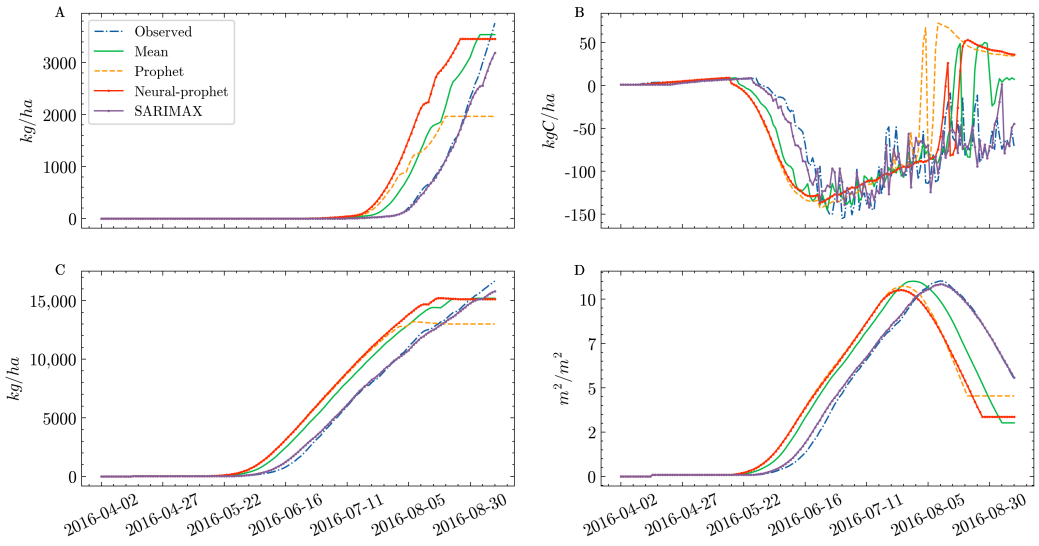

Figure 6. Season dynamics of yield (**A**), NEE (**B**), aboveground biomass (**C**), and LAI (**D**) under different weather scenarios for soybean.

The forecast accuracy indices are provided in Table 4. In the case of the yield prediction, the grouping by mean variant obtained the DTW score of 403.71. However, the SARIMAX model shows the R^2 of 0.98, which is higher than 0.80 for the averaging variant. SARIMAX also obtained the highest quality scores in the case of the aboveground biomass dynamics, LAI, and NEE. It is also worth noting that all models were worse at predicting NEE dynamics than all the other parameters.

Table 4. Soybean dynamics forecast quality indices.

	Crop Yield			Biomass			LAI			NEE		
	RRMSE	**R^2**	**DTW**	**RRMSE**	**R^2**	**DTW**	**RRMSE**	**R^2**	**DTW**	**RRMSE**	**R^2**	**DTW**
Baseline	0.44	0.80	**403.71**	0.19	0.96	3187.21	0.32	0.89	7.68	0.51	0.73	326.01
SARIMAX	**0.12**	**0.98**	969.97	**0.04**	**0.99**	**1682.33**	**0.04**	**0.99**	**0.58**	**0.35**	**0.88**	**184.21**
Prophet	0.49	0.75	4194.16	0.27	0.92	11,062.58	0.41	0.83	4.47	0.9	0.19	280.57
Neural Prophet	0.74	0.40	503.84	0.27	0.92	3435.52	0.43	0.81	8.87	0.64	0.58	280.47

Bold indicates best model performance.

The growth curves and dynamics of the key biophysical parameters for sugar beet are provided in Figure 7. In the case of sugar beet, unlike soybean, the models overestimate yields compared to the observations. The grouping by mean variant gives earlier crop maturation and a significant overestimation of yield. The Prophet, Neural Prophet, and SARIMAX models also overestimate yields. In contrast to yield, the aboveground biomass dynamics are better estimated by the different models. It is worth noting that unlike yield, the aboveground biomass values are underestimated by SARIMAX. It is also worth mentioning that, unlike soybean, sugar beet yields show less correlation with the aboveground biomass.

The dynamics of ecosystem carbon exchange in the case of real-life observations exhibit a rather high amplitude over the vegetation season. Prophet and Neural Prophet give smoother NEE dynamics compared to the real observations, which was the same for soybean. In turn, the SARIMAX model gives an amplitude that is closer to the observed NEE dynamics.

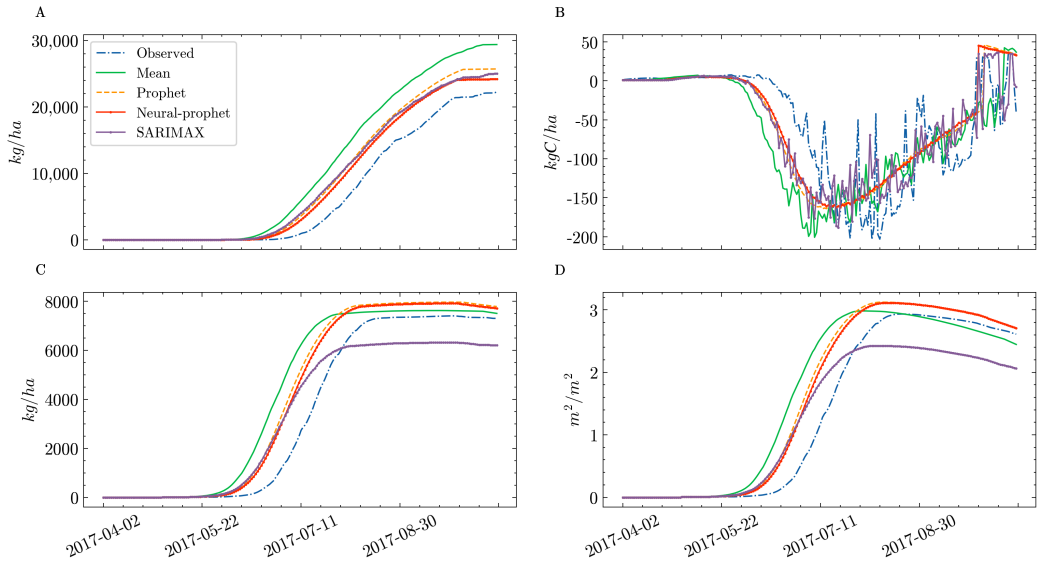

Figure 7. Season dynamics of yield (**A**), NEE (**B**), aboveground biomass (**C**), and LAI (**D**) under different weather scenarios for sugar beet.

In the case of leaf area index, the following pattern was observed: all variants produced a forecast that was biased towards the start of the season. In the second half of the season, the Prophet and Neural Prophet models predict the LAI dynamics that replicate the observed dynamics quite well. The SARIMAX model, in its turn, underestimates the LAI values in the second half of the season by about 0.5 units. The Neural Prophet model gives the closest value to the observed estimate of yield with the total value of 24,187 kg/ha, and the worst forecast is produced by the weather averaging approach with the value of 29,395 kg/ha, which is 32% higher than the observed yield. In contrast to soybean, for sugar beet, all forecasts overestimated the total yield. We should also note a rather large variation in the totals of the precipitation values obtained by different models, starting from 291 mm in the observations to 556 mm for the Prophet model.

The forecast accuracy indices for sugar beet are provided in Table 5. The Neural Prophet model revealed the highest quality index values in the case of the yield forecast, in particular, R^2 = 0.95. The Neural Prophet model also obtained the best R^2 values for all other biophysical indicators compared to other models.

Table 5. Sugar beet dynamics forecast quality indices.

	Crop Yield			**Biomass**			**LAI**			**NEE**		
	RRMSE	**R^2**	**DTW**	**RRMSE**	**R^2**	**DTW**	**RRMSE**	**R^2**	**DTW**	**RRMSE**	**R^2**	**DTW**
Baseline	0.50	0.72	44,301	0.34	0.88	**1786**	0.35	0.87	**0.59**	0.18	0.57	284
SARIMAX	0.27	0.92	14,986	0.24	0.94	9664	0.27	0.92	4.40	**0.05**	0.71	**241**
Prophet	0.29	0.91	20,650	0.25	0.93	4876	0.25	0.93	1.07	0.53	0.72	342
Neural Prophet	**0.22**	**0.95**	**11,019**	**0.22**	**0.95**	4313	**0.22**	**0.4**	1.02	0.50	**0.75**	343

Bold indicates best model performance.

3.3. Soil Dynamics Modeling

The soil conditions in the plough horizon obtained in the soybean simulations are presented in Figure 8. As can be seen on the plot, most models gave poor quality weather forecasts when reproducing soil moisture. Prophet and Neural Prophet underpredicted soil moisture compared to the real weather observations. Using SARIMAX and the grouping by mean model allowed soil moisture values to be obtained that were closer to the actual observation, although a certain underestimation can also be seen. In the summer season, a sharp increase in soil humidity can be seen in the case of the real-life observations with peaks exceeding 0.35, while the predicted peak values in the corresponding period are not higher than about 0.3.

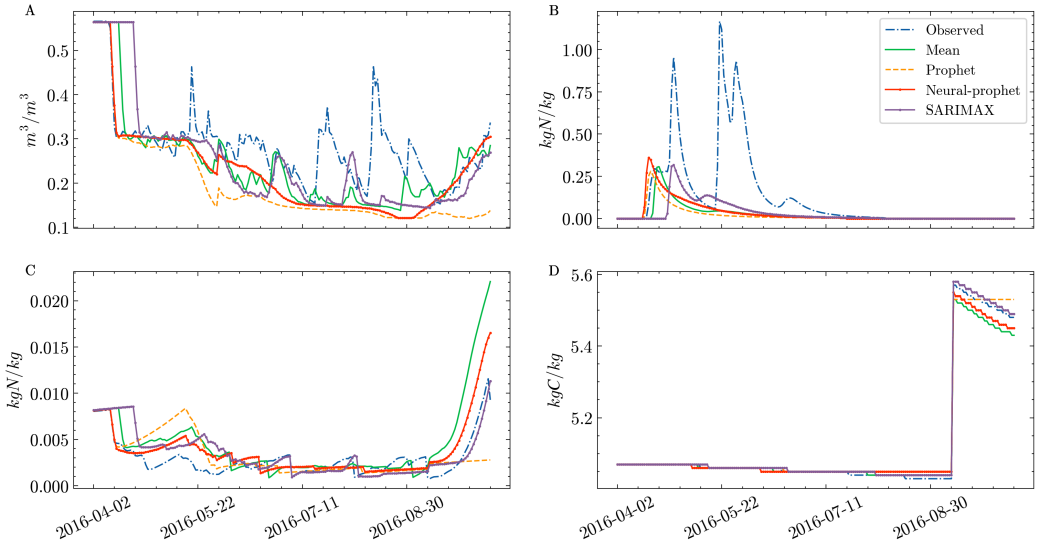

Figure 8. Season dynamics of soil humidity (**A**), nitrate leaching (**B**), nitrate content (**C**), and organic matter carbon content (**D**) under different weather scenarios for soybean.

The nitrate leaching dynamics reveal a pattern similar to soil humidity. None of the weather forecasting methods obtained the same high leaching values as in the real observations. All the models and grouping by mean produce similar dynamics and underestimate the maximum leaching values, from 1.0–1.2 to 0.3–0.4 kgN/ha. It is also worth noting that the Prophet and Neural Prophet models produce a peak of nitrogen leaching at the start of the season, while two peaks can be distinguished in the real weather observations, one of them occurring in early May and the other in early June. This may be due to the presence of high precipitation, as can be seen in the Figure A1. SARIMAX predicts the first peak in early May, similar to the real observations, but the second peak in early June is not predicted.

The amount of precipitation and its distribution over the season play a significant role in forming the dynamics of soil moisture and nitrate leaching. As mentioned above, precipitation is one of the parameters that are challenging to predict because of its stochasticity, which in turn impacts the possibility of modeling those soil conditions for which precipitation is the main affecting factor. For this reason, smoothed precipitation predictions underestimate the moisture levels and may lead to the underprediction of nitrate leaching.

The nitrate content dynamics do not show such pronounced fluctuations during the season as moisture and nitrate leaching. For both model forecasts and real weather observations, the nitrate content level appears to be very low throughout the season and increases only after crop harvesting in early September. The closest to the observed dynamics was achieved with the SARIMAX model. As can be seen on the plot, grouping by mean and the Neural Prophet model overestimated the nitrate content at the end of the season, while the Prophet model gave nitrate content levels that were lower than the observed ones. The organic matter content showed a dynamic that is similar to the nitrate content. The plot shows an increase in the organic matter content after harvesting, which is associated with the residual plant parts left in the soil horizon. Similar to the case with the nitrates dynamic, SARIMAX allowed us to obtain a forecast closely replicating the real-life observations.

Table 6 compares the quality indices for the soil condition forecasting models in the case of soybean. As the table shows, the RRMSE and R^2 for soil humidity and nitrate leaching are worse than those for the organic matter and nitrate content dynamics, which can be also seen in Figure 8. The same tendency was observed for all the weather scenario modeling methods in this study. When the models were compared to each other, the SARIMAX model achieved the best quality metrics values for all soil parameters.

Table 6. Soil parameters dynamics forecast quality indices for soybean.

	Soil Organic Carbon			Soil Moisture			Nitrate Leaching			Nitrate Content		
	RRMSE	R^2	**DTW**	**RRMSE**	R^2	**DTW**	**RRMSE**	R^2	**DTW**	**RRMSE**	R^2	**DTW**
Baseline	0.18	0.96	0.003	0.69	0.46	1.590	0.92	0.07	2.27	0.82	0.32	0.180
SARIMAX	**0.05**	**0.99**	**0.001**	**0.55**	**0.68**	**0.510**	**0.86**	**0.19**	2.24	**0.25**	**0.93**	**0.030**
Prophet	0.28	0.92	0.004	0.71	0.31	1.750	0.95	0.00	2.35	0.73	0.41	0.170
Neural Prophet	0.13	0.98	0.003	0.67	0.46	1.710	0.92	0.08	**2.12**	0.76	0.40	0.180

Bold indicates best model performance.

The soil conditions in the plough horizon obtained in the sugar beet simulations are presented in Figure 9. For sugar beet soil moisture, significant differences between the observations and model forecasts were also obtained.

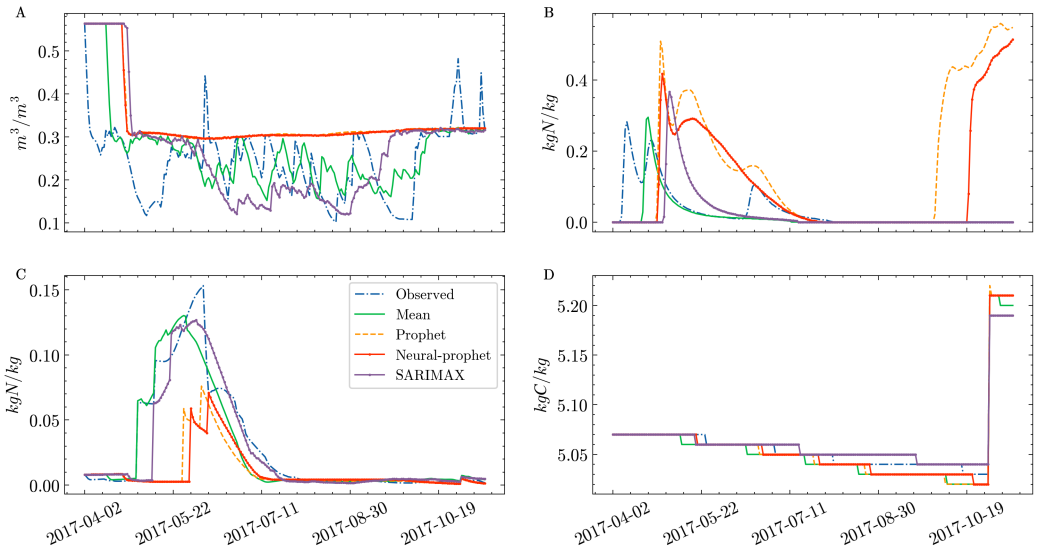

Figure 9. Season dynamics of soil humidity (**A**), nitrate leaching (**B**), nitrate content (**C**), and organic matter carbon content (**D**) under different weather scenarios for sugar beet.

However, unlike the case with soybean, the sugar beet models tended to overestimate soil moisture levels. The Prophet and Neural Prophet models allowed smoothed moisture dynamics with values of about 0.3 to be obtained, while the real observations showed more pronounced fluctuations and ranged between 0.1 and 0.3 on average. The SARIMAX model and the grouping by mean variant showed dynamics close to the observed one. Nitrate leaching shows a significant increase at the start of the season and then a gradual decrease, which is associated with the use of nitrogen fertilizers, which are taken into account in the modeling. The grouping by mean variant produced a dynamic that was closest to the real-life observations. Prophet and Neural Prophet showed some anomalous increases at the end of the season, which may be mentioned as shortcomings of these forecasts. The nitrate content dynamics showed a sharp peak in the first half of the season after nitrogen fertilizers were applied. The SARIMAX and the grouping by mean models produced forecasts that were quite close to the real content, while the Prophet and Neural Prophet models underestimated the content by about two times. The dynamics of the organic matter content are similar to the soybean case and show a growth after crop harvesting resulting from the plant residues remaining in the plough horizon.

The best quality soil condition forecasts for sugar beet were obtained using weather averaging over several years (Table 7). Neural Prophet obtained the best quality indices for the soil organic matter forecast only, although the differences between this model and the averaging variant were not significant. We should also note the poor quality of the nitrogen leaching forecasts obtained with the models, with only the grouping by mean variant obtaining a positive R^2 value equal to 0.34. It can be assumed that the quality of the soil dynamics prediction depends on the quality of the crop dynamics prediction.

Table 7. Soil parameters dynamics forecast quality indices for sugar beet.

	Soil Organic Carbon			Soil Moisture			Nitrate Leaching			Nitrate Content		
	RRMSE	**R^2**	**DTW**	**RRMSE**	**R^2**	**DTW**	**RRMSE**	**R^2**	**DTW**	**RRMSE**	**R^2**	**DTW**
Baseline	0.17	0.97	0.0007	**0.54**	**0.7**	**0.52**	**0.79**	**0.34**	**0.26**	**0.29**	**0.91**	**0.06**
SARIMAX	0.24	0.94	0.0016	0.71	0.5	0.75	1.29	–0.69	0.29	0.35	0.87	0.08
Prophet	0.16	0.97	0.0007	0.74	0.4	1.82	1.8	–25.25	4.31	0.81	0.25	0.25
Neural Prophet	**0.13**	**0.98**	**0.0006**	0.61	0.57	1.38	2.02	–9.56	1.65	0.86	0.16	0.28

Bold indicates best model performance.

As part of our research, we generalized machine learning and statistical analysis methods for working with climate data. We identified which variables can be accurately predicted using these methods and which contain uncertainty that cannot be predicted. In addition, based on the identified results, it is possible to conclude the contribution of individual variables to subsequent modeling using simulation methods.

We hope that our research might be helpful for the development of methods and techniques for the cheap generation of surrogate weather scenarios and risk estimations. However, such research might require an additional scale of computations related to one more mathematical model for risks. Nevertheless, our research shows the advantages and limitations of popular time series forecasting methods for different weather and soil parameters. We see that some parameters can be predicted accurately with our approach, and some require another approach to be found.

One of the main limitations of most seasonal forecast approaches is the complexity of modeling extreme events. This drawback has a powerful effect on the ability to predict precipitation [52]. However, this is important for modeling nitrogen leaching for soybeans. As can be seen, for observed weather, there are several peaks of nitrogen leaching due to extreme precipitation. At the same time, only one peak is observed for the model variants since the models did not predict extreme precipitation. Thus, there may be an underestimation of essential factors, such as nitrogen leaching, which directly depend on the possibility of predicting extreme events. It is important to note that within the framework of the presented study, meteorological parameters were evaluated as independent time series data. In subsequent studies, it makes sense to consider the possibility of conducting a correlation analysis between the abovementioned variables.

Future studies could explore the reasons behind the underestimation of crop yield, explicitly investigating the role of poor precipitation forecasting in model predictions. Additionally, using transformer-based methods that are popular in NLP tasks can improve the quality of the forecast and the use of hybrid approaches based on statistical weather generators and machine learning models [53,54].

4. Conclusions

The numerical experiments conducted in the present work demonstrated the possibility of using machine learning methods for time series forecasting to predict weather scenarios, which could be used in the simulation modeling of agrosystems. The SARIMAX and Neural Prophet models allowed soybean and sugar beet yield estimates to be obtained that were rather close to the observed weather conditions. However, the time series forecasting methods used smooth the series of weather observations, which seems less critical in the case of predicting soil nutrient scenarios, but can make modeling soil moisture dynamics quite challenging since dynamics smoothing produced by these models does not allow extreme events (e.g., heavy rainfall) to be taken into account. Currently, the effectiveness of long-term weather forecasts is limited due to the high level of uncertainty. However, as forecasting technology continues to advance and become more accurate, it is expected that agricultural systems will benefit from improved forecasting performance.

Author Contributions: Conceptualization, I.M., S.M. and M.G.; funding acquisition, M.G.; methodology, S.M. and M.G.; project administration, S.M. and M.G.; resources, M.G.; software, I.M.; supervision, S.M.; visualization, M.G.; writing—original draft, I.M.; writing—review and editing, S.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Russian Science Foundation (grant no. 23-21-00336).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Input data used for the research and source code are freely available from the GitHub—https://github.com/mirpulatov/monica_modeling (accessed on 18 August 2023).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MSE	Mean Squared Error
RRMSE	Relative Root Mean Squared Error
R^2	R-squared
ARIMA	Autoregressive Integrated Moving Average
SARIMAX	Seasonal Autoregressive Integrated Moving Average with exogenous regressors
DTW	Dynamic time warping
LAI	Leaf area index
NEE	Net ecosystem carbon exchange

Appendix A

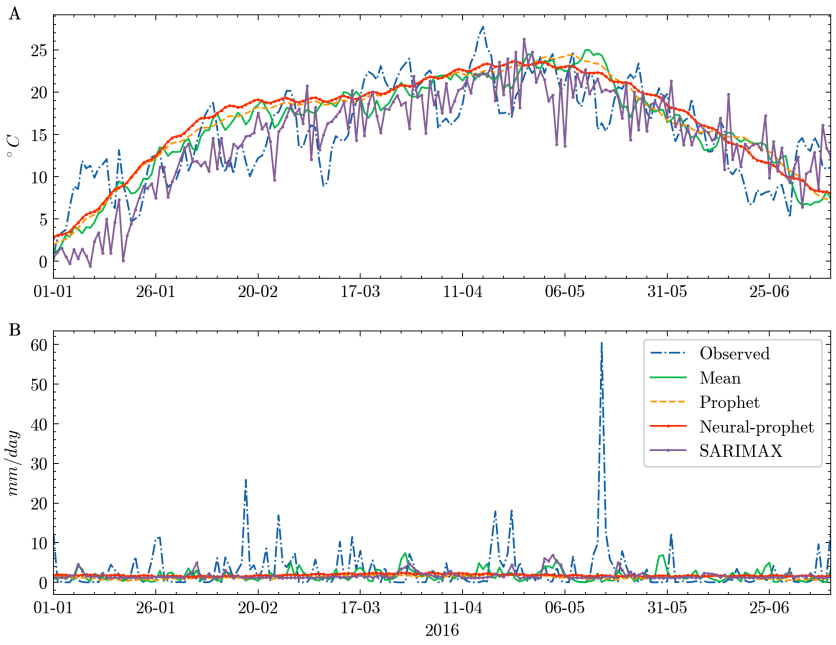

Figure A1. Weather scenario forecasting for average temperature (**A**) and precipitation (**B**) for 2016 using different methods.

References

1. Howden, S.M.; Soussana, J.F.; Tubiello, F.N.; Chhetri, N.; Dunlop, M.; Meinke, H. Adapting agriculture to climate change. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 19691–19696. [CrossRef] [PubMed]
2. Liu, K.; Harrison, M.T.; Yan, H.; Liu, D.L.; Meinke, H.; Hoogenboom, G.; Wang, B.; Peng, B.; Guan, K.; Jaegermeyr, J.; et al. Silver lining to a climate crisis in multiple prospects for alleviating crop waterlogging under future climates. *Nat. Commun.* **2023**, *14*, 765. [CrossRef] [PubMed]
3. Lal, R. Food security in a changing climate. *Ecohydrol. Hydrobiol.* **2013**, *13*, 8–21. [CrossRef]

4. Bassu, S.; Brisson, N.; Durand, J.L.; Boote, K.; Lizaso, J.; Jones, J.W.; Rosenzweig, C.; Ruane, A.C.; Adam, M.; Baron, C.; et al. How do various maize crop models vary in their responses to climate change factors? *Glob. Chang. Biol.* **2014**, *20*, 2301–2320. [CrossRef]
5. Peng, B.; Guan, K.; Tang, J.; Ainsworth, E.A.; Asseng, S.; Bernacchi, C.J.; Cooper, M.; Delucia, E.H.; Elliott, J.W.; Ewert, F.; et al. Towards a multiscale crop modelling framework for climate change adaptation assessment. *Nat. Plants* **2020**, *6*, 338–348.
6. Petrovskaia, A.; Ryzhakov, G.; Oseledets, I. Optimal soil sampling design based on the maxvol algorithm. *Geoderma* **2021**, *402*, 115362. [CrossRef]
7. Webber, H.; Hoffmann, M.; Rezaei, E.E. Crop models as tools for agroclimatology. *Agroclimatol. Link. Agric. Clim.* **2020**, *60*, 519–546.
8. Rurinda, J.; Zingore, S.; Jibrin, J.M.; Balemi, T.; Masuki, K.; Andersson, J.A.; Pampolino, M.F.; Mohammed, I.; Mutegi, J.; Kamara, A.Y.; et al. Science-based decision support for formulating crop fertilizer recommendations in sub-Saharan Africa. *Agric. Syst.* **2020**, *180*, 102790. [CrossRef] [PubMed]
9. Jones, J.W.; Hoogenboom, G.; Porter, C.H.; Boote, K.J.; Batchelor, W.D.; Hunt, L.; Wilkens, P.W.; Singh, U.; Gijsman, A.J.; Ritchie, J.T. The DSSAT cropping system model. *Eur. J. Agron.* **2003**, *18*, 235–265.
10. Holzworth, D.P.; Huth, N.I.; deVoil, P.G.; Zurcher, E.J.; Herrmann, N.I.; McLean, G.; Chenu, K.; van Oosterom, E.J.; Snow, V.; Murphy, C.; et al. APSIM–evolution towards a new generation of agricultural systems simulation. *Environ. Model. Softw.* **2014**, *62*, 327–350.
11. Brisson, N.; Gary, C.; Justes, E.; Roche, R.; Mary, B.; Ripoche, D.; Zimmer, D.; Sierra, J.; Bertuzzi, P.; Burger, P.; et al. An overview of the crop model STICS. *Eur. J. Agron.* **2003**, *18*, 309–332. [CrossRef]
12. Badenko, V.; Terleev, V.; Topaj, A. AGROTOOL software as an intellectual core of decision support systems in computer aided agriculture. *Appl. Mech. Mater.* **2014**, *635*, 1688–1691. [CrossRef]
13. Nendel, C. MONICA: A simulation model for nitrogen and carbon dynamics in agro-ecosystems. In *Novel Measurement and Assessment Tools for Monitoring and Management of Land and Water Resources in Agricultural Landscapes of Central Asia*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 389–405.
14. Porter, C.H.; Jones, J.W.; Adiku, S.; Gijsman, A.J.; Gargiulo, O.; Naab, J. Modeling organic carbon and carbon-mediated soil processes in DSSAT v4. 5. *Oper. Res.* **2010**, *10*, 247–278.
15. Liu, H.; Yang, J.; Tan, C.; Drury, C.; Reynolds, W.; Zhang, T.; Bai, Y.; Jin, J.; He, P.; Hoogenboom, G. Simulating water content, crop yield and nitrate-N loss under free and controlled tile drainage with subsurface irrigation using the DSSAT model. *Agric. Water Manag.* **2011**, *98*, 1105–1111. [CrossRef]
16. Vogeler, I.; Cichota, R.; Langer, S.; Thomas, S.; Ekanayake, D.; Werner, A. Simulating water and nitrogen runoff with APSIM. *Soil Tillage Res.* **2023**, *227*, 105593. [CrossRef]
17. Poluektov, R.A.; Terleev, V.V. Crop simulation model of the second and the third productivity levels. In Proceedings of the Modelling Water and Nutrient Dynamics in Soil–Crop Systems: Proceedings of the Workshop on "Modelling Water and Nutrient Dynamics in Soil–Crop Systems", Müncheberg, Germany, 14–16 June 2004; Springer: Berlin/Heidelberg, Germany, 2007; pp. 75–89.
18. Semenov, M.A. Simulation of extreme weather events by a stochastic weather generator. *Clim. Res.* **2008**, *35*, 203–212. [CrossRef]
19. Ma, D.; Jing, Q.; Xu, Y.P.; Cannon, A.J.; Dong, T.; Semenov, M.A.; Qian, B. Using ensemble-mean climate scenarios for future crop yield projections: A stochastic weather generator approach. *Clim. Res.* **2021**, *83*, 161–171. [CrossRef]
20. Confalonieri, R. Combining a weather generator and a standard sensitivity analysis method to quantify the relevance of weather variables on agrometeorological models outputs. *Theor. Appl. Climatol.* **2012**, *108*, 19–30. [CrossRef]
21. Sidhu, R.K.; Kumar, R.; Rana, P.S. Long short-term memory neural network-based multi-level model for smart irrigation. *Mod. Phys. Lett. B* **2020**, *34*, 2050418. [CrossRef]
22. Schultz, M.G.; Betancourt, C.; Gong, B.; Kleinert, F.; Langguth, M.; Leufen, L.H.; Mozaffari, A.; Stadtler, S. Can deep learning beat numerical weather prediction? *Philos. Trans. R. Soc. A* **2021**, *379*, 20200097. [CrossRef]
23. Weyn, J.A.; Durran, D.R.; Caruana, R.; Cresswell-Clay, N. Sub-seasonal forecasting with a large ensemble of deep-learning weather prediction models. *J. Adv. Model. Earth Syst.* **2021**, *13*, e2021MS002502. [CrossRef]
24. Eyring, V.; Bony, S.; Meehl, G.A.; Senior, C.A.; Stevens, B.; Stouffer, R.J.; Taylor, K.E. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* **2016**, *9*, 1937–1958. [CrossRef]
25. Kheir, A.M.; Elnashar, A.; Mosad, A.; Govind, A. An improved deep learning procedure for statistical downscaling of climate data. *Heliyon* **2023**, *9*, e18200. [CrossRef] [PubMed]
26. Van Straaten, C.; Whan, K.; Coumou, D.; Van den Hurk, B.; Schmeits, M. Using explainable machine learning forecasts to discover subseasonal drivers of high summer temperatures in western and central Europe. *Mon. Weather. Rev.* **2022**, *150*, 1115–1134. [CrossRef]
27. Weirich-Benet, E.; Pyrina, M.; Jiménez-Esteve, B.; Fraenkel, E.; Cohen, J.; Domeisen, D.I. Subseasonal Prediction of Central European Summer Heatwaves with Linear and Random Forest Machine Learning Models. *Artif. Intell. Earth Syst.* **2023**, *2*, e220038. [CrossRef]
28. Slater, L.; Arnal, L.; Boucher, M.A.; Chang, A.Y.Y.; Moulds, S.; Murphy, C.; Nearing, G.; Shalev, G.; Shen, C.; Speight, L.; et al. Hybrid forecasting: Using statistics and machine learning to integrate predictions from dynamical models. *Hydrol. Earth Syst. Sci. Discuss* 2022, *preprint*. [CrossRef]

29. Slater, L.J.; Arnal, L.; Boucher, M.A.; Chang, A.Y.Y.; Moulds, S.; Murphy, C.; Nearing, G.; Shalev, G.; Shen, C.; Speight, L.; et al. Hybrid forecasting: Blending climate predictions with AI models. *Hydrol. Earth Syst. Sci.* **2023**, *27*, 1865–1889. [CrossRef]
30. Micheli, E.; Schad, P.; Spaargaren, O.; Dent, D. World reference base for soil resources. In *World Soil Resources Reports*, FAO: Rome, Italy, 2006.
31. Bai, J.; Chen, X.; Dobermann, A.; Yang, H.; Cassman, K.G.; Zhang, F. Evaluation of NASA satellite-and model-derived weather data for simulation of maize yield potential in China. *Agron. J.* **2010**, *102*, 9–16. [CrossRef]
32. Van Wart, J.; Grassini, P.; Cassman, K.G. Impact of derived global weather data on simulated crop yields. *Glob. Chang. Biol.* **2013**, *19*, 3822–3834. [CrossRef]
33. Savary, S.; Nelson, A.; Willocquet, L.; Pangga, I.; Aunario, J. Modeling and mapping potential epidemics of rice diseases globally. *Crop Prot.* **2012**, *34*, 6–17. [CrossRef]
34. Gasanov, M.; Merkulov, D.; Nikitin, A.; Matveev, S.; Stasenko, N.; Petrovskaia, A.; Pukalchik, M.; Oseledets, I. A New Multi-objective Approach to Optimize Irrigation Using a Crop Simulation Model and Weather History. In Proceedings of the Computational Science–ICCS 2021: 21st International Conference, Krakow, Poland, 16–18 June 2021; Proceedings Part IV; Springer: Berlin/Heidelberg, Germany, 2021; pp. 75–88.
35. Duarte, Y.C.; Sentelhas, P.C. NASA/POWER and DailyGridded weather datasets—How good they are for estimating maize yields in Brazil? *Int. J. Biometeorol.* **2020**, *64*, 319–329. [CrossRef]
36. Rodrigues, G.C.; Braga, R.P. Estimation of daily reference evapotranspiration from NASA POWER reanalysis products in a hot summer mediterranean climate. *Agronomy* **2021**, *11*, 2077. [CrossRef]
37. Pawson, S.; Stolarski, R.S.; Douglass, A.R.; Newman, P.A.; Nielsen, J.E.; Frith, S.M.; Gupta, M.L. Goddard Earth Observing System chemistry-climate model simulations of stratospheric ozone-temperature coupling between 1950 and 2005. *J. Geophys. Res. Atmos.* **2008**, *113*. [CrossRef]
38. Box, G.E.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. *Time Series Analysis: Forecasting and Control*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
39. Arunraj, N.S.; Ahrens, D.; Fernandes, M. Application of SARIMAX model to forecast daily sales in food retail industry. *Int. J. Oper. Res. Inf. Syst. (IJORIS)* **2016**, *7*, 1–21. [CrossRef]
40. Taylor, S.J.; Letham, B. Forecasting at scale. *Am. Stat.* **2018**, *72*, 37–45. [CrossRef]
41. Triebe, O.; Hewamalage, H.; Pilyugina, P.; Laptev, N.; Bergmeir, C.; Rajagopal, R. Neuralprophet: Explainable forecasting at scale. *arXiv* **2021**, arXiv:2111.15397.
42. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
43. Fouladgar, N.; Alirezaie, M.; Främling, K. Metrics and Evaluations of Time Series Explanations: An Application in Affect Computing. *IEEE Access* **2022**, *10*, 23995–24009. [CrossRef]
44. Senin, P. Dynamic time warping algorithm review. *Inf. Comput. Sci.* **2008**, *855*, 40.
45. Kersebaum, K.C. Modelling nitrogen dynamics in soil–crop systems with HERMES. In Proceedings of the Modelling Water and Nutrient Dynamics in Soil–Crop Systems: Proceedings of the Workshop on "Modelling water and Nutrient Dynamics in soil–crop Systems", Müncheberg, Germany, 14–16 June 2004; Springer: Berlin/Heidelberg, Germany, 2007; pp. 147–160.
46. Mirschel, W.; Wenkel, K.O. Modelling soil–crop interactions with AGROSIM model family. In Proceedings of the Modelling Water and Nutrient Dynamics in Soil–Crop Systems: Proceedings of the Workshop on "Modelling Water and Nutrient Dynamics in Soil–Crop Systems", Müncheberg, Germany, 14–16 June 2004; Springer: Berlin/Heidelberg, Germany, 2007; pp. 59–73.
47. Abrahamsen, P.; Hansen, S. Daisy: An open soil-crop-atmosphere system model. *Environ. Model. Softw.* **2000**, *15*, 313–330. [CrossRef]
48. Nendel, C.; Berg, M.; Kersebaum, K.C.; Mirschel, W.; Specka, X.; Wegehenkel, M.; Wenkel, K.; Wieland, R. The MONICA model: Testing predictability for crop growth, soil moisture and nitrogen dynamics. *Ecol. Model.* **2011**, *222*, 1614–1625. [CrossRef]
49. Rötter, R.P.; Palosuo, T.; Kersebaum, K.C.; Angulo, C.; Bindi, M.; Ewert, F.; Ferrise, R.; Hlavinka, P.; Moriondo, M.; Nendel, C.; et al. Simulation of spring barley yield in different climatic zones of Northern and Central Europe: A comparison of nine crop models. *Field Crop. Res.* **2012**, *133*, 23–36. [CrossRef]
50. Specka, X.; Nendel, C.; Wieland, R. Analysing the parameter sensitivity of the agro-ecosystem model MONICA for different crops. *Eur. J. Agron.* **2015**, *71*, 73–87. [CrossRef]
51. Gasanov, M.; Petrovskaia, A.; Nikitin, A.; Matveev, S.; Tregubova, P.; Pukalchik, M.; Oseledets, I. Sensitivity analysis of soil parameters in crop model supported with high-throughput computing. In *International Conference on Computational Science*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 731–741.
52. Qian, Q.; Jia, X.; Lin, H.; Zhang, R. Seasonal forecast of nonmonsoonal winter precipitation over the Eurasian continent using machine-learning models. *J. Clim.* **2021**, *34*, 7113–7129.
53. Shin, J.Y.; Kim, K.R.; Ha, J.C. Seasonal forecasting of daily mean air temperatures using a coupled global climate model and machine learning algorithm for field-scale agricultural management. *Agric. For. Meteorol.* **2020**, *281*, 107858. [CrossRef]
54. Gao, Z.; Shi, X.; Wang, H.; Zhu, Y.; Wang, Y.B.; Li, M.; Yeung, D.Y. Earthformer: Exploring space-time transformers for earth system forecasting. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 25390–25403.