

## Article

# A Forecast Model Applied to Monitor Crops Dynamics Using Vegetation Indices (NDVI)

Francisco Carreño-Conde <sup>1,2,\*</sup> , Ana Elizabeth Sipols <sup>3</sup> , Clara Simón de Blas <sup>4,5</sup>  and David Mostaza-Colado <sup>6</sup> 

<sup>1</sup> Department of Biology and Geology, Physics and Inorganic Chemistry, Higher School of Experimental Sciences and Technology, Rey Juan Carlos University, c/Tulipán s/n, 28933 Móstoles, Spain

<sup>2</sup> IMDEA Water Institute, Avenida Punto Com 2, Parque Científico Tecnológico de la Universidad de Alcalá, 28805 Alcalá de Henares, Spain

<sup>3</sup> Department of Computer science and Statistics, Engineering Computer Science School, Rey Juan Carlos University, c/Tulipán s/n, 28933 Móstoles, Spain; anaelizabeth.garcia@urjc.es

<sup>4</sup> Department of Applied Mathematics, Material Science and Engineering and Electronic Technology, Higher School of Experimental Sciences and Technology, Rey Juan Carlos University, c/Tulipán s/n, 28933 Móstoles, Spain; clara.simon@urjc.es

<sup>5</sup> Instituto Universitario de Evaluación Sanitaria, Facultad de Medicina Pabellón 4 Planta 1, UCM, Ciudad Universitaria, 28040 Madrid, Spain

<sup>6</sup> Agro-Environmental Research Department, Madrid Institute of Rural, Agricultural and Food Research and Development (IMIDRA), Ctra. A2, Km.38.200, 28805 Alcalá de Henares, Spain; david.mostaza@madrid.org

\* Correspondence: francisco.carreno@urjc.es; Tel.: +34-91-488-83-05

**Abstract:** Vegetation dynamics is very sensitive to environmental changes, particularly in arid zones where climate change is more prominent. Therefore, it is very important to investigate the response of this dynamics to those changes and understand its evolution according to different climatic factors. Remote sensing techniques provide an effective system to monitor vegetation dynamics on multiple scales using vegetation indices (VI), calculated from remote sensing reflectance measurements in the visible and infrared regions of the electromagnetic spectrum. In this study, we use the normalized difference vegetation index (NDVI), provided from the MOD13Q1 V006 at 250 m spatial resolution product derived from the MODIS sensor. NDVI is frequent in studies related to vegetation mapping, crop state indicator, biomass estimator, drought monitoring and evapotranspiration. In this paper, we use a combination of forecasts to perform time series models and predict NDVI time series derived from optical remote sensing data. The proposed ensemble is constructed using forecasting models based on time series analysis, such as Double Exponential Smoothing and autoregressive integrated moving average with explanatory variables for a better prediction performance. The method is validated using different maize plots and one olive plot. The results after combining different models show the positive influence of several weather measures, namely, temperature, precipitation, humidity and radiation.

**Keywords:** forecast model; NDVI; crops; remote sensing; MODIS; weather variables; time series analysis; ARIMAX



**Citation:** Carreño-Conde, F.; Sipols, A.E.; de Blas, C.S.; Mostaza-Colado, D. A Forecast Model Applied to Monitor Crops Dynamics Using Vegetation Indices (NDVI). *Appl. Sci.* **2021**, *11*, 1859. <https://doi.org/10.3390/app11041859>

Academic Editor: Lorena Parra

Received: 15 January 2021

Accepted: 16 February 2021

Published: 20 February 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Vegetation dynamic is very sensitive to environmental changes, particularly in arid zones where climate change is more prominent. Therefore, it is very important to investigate the response of this dynamics to those changes and understand its evolution according to different climatic factors. Pointed out as a cumulative effect on vegetation in combination with environmental changes that generates a temporary delay in the response of plants, this must be considered in order to understand the variations in vegetation and predict its changing characteristics under future climate changes [1–5].

According to this, several studies claim that the main factors that determine the relationship between vegetation and climate change are temperature and precipitation [6].

Therefore, the water content of the unsaturated zone determines the optimal development of vegetation. The availability of water in the soil over a period of time (year, month and day) is conditioned by the inflows of water, the effective precipitation reaching the soil and the water reserves accumulated in the soil itself [7,8]. However, the loss of moisture in the soil occurs as a consequence of the direct evaporation of water from the soil and the transpiration of the plants, which are direct functions of the ambient temperature and soil temperature [9–12].

Remote sensing techniques provide an effective system to monitor vegetation dynamics on multiple scales using vegetation indices (IV) that are based on the difference of near infrared (NIR) and red visible (R) spectral reflectance [13]. The normalized difference vegetation index (NDVI) is the most widely used because it quantifies vegetation health by measuring the difference between near infrared (strongly reflected) and red light (absorbed). This index (Equation (1)) is calculated from remote sensing measurements in the visible and infrared bands as:

$$NDVI = \frac{(\sigma_{NIR} - \sigma_R)}{(\sigma_{NIR} + \sigma_R)} \quad (1)$$

where  $\rho_{NIR}$  is the surface reflectance in the near infrared region and  $\rho_R$  is the surface reflectance in the red visible region of the electromagnetic spectrum. The NDVI of a vegetated area tends toward positive values, whereas water and urban areas are represented by near zero or negative values [14]. This relationship between NDVI with vegetation health is well known and used, among other studies, for vegetation mapping, as a crop state indicator, as a biomass estimator and for drought or evapotranspiration monitoring [15–22].

In accordance with the above time series, vegetation indices with high spatial resolution and high temporal frequency are important to describe and analyze the spatial patterns for vegetation monitoring and management [23]. Additionally, important issues, such as the prediction of vegetation response induced for climatic factors, are based on forecasting models of NDVI [24].

NDVI series have been studied in different ways applying several analyses. Han et al. (2010) proposed ARIMA models to simulate the vegetation temperature condition index (VTCI) series, and forecast their changes in the future [25]. In addition, the (2010) MODIS LAI (Leaf Area Index) model proposed by Jiang et al. applies three models that can characterize the non-stationary time series data and predict the future values, including Dynamic Harmonics Regression (DHR), Seasonal-Trend Decomposition Procedure based on Loess (STL) and seasonal autoregressive integrated moving average (SARIMA) [26].

Fernández-Manso et al. (2011) developed a method for forecasting the short-term response of forest vegetation on the basis of an autoregressive integrated moving average (ARIMA) [27]. Gonçalves et al. (2012) applied different statistical forecasting methods to an agroclimatic index (the water requirement satisfaction index: WRSI) and the sugarcane spectral response (normalized difference vegetation index: NDVI) registered on National Oceanic and Atmospheric Administration Advanced Very High Resolution Radiometer (NOAA-AVHRR) satellite images [28]. Rhee and Im (2017) tested drought forecast using long-term forecast data and compared it with using just climatological data to obtain drought forecast information in uncalibrated areas. Additionally, they used machine learning techniques of long-term forecast data or climatological data combined with remote sensing data to compare the performance with Kriging spatial interpolation [29]. Reddy and Prasad (2018) predicted the vegetation dynamics using MODIS NDVI time series data sets, a long short term memory network and an artificial neural network [30]. Most recently, Ba et al. (2020) and Li et al. (2017) approached the study of NDVI series using multi-scaling analysis [31,32].

The objective of this paper is to offer an algorithm for forecasting vegetation indices (NDVI) time series derived from optical remote sensing data. Specifically, we use NDVI data derived from a Moderate Resolution Imaging Spectroradiometer sensor (MODIS) for different types of crops across seasons during the February 2013–December 2017 period.

This main objective can help to resolve the following questions: (1) Does it make sense to combine forecasting model temporal series to obtain a more accurate prediction model? (2) What weather parameters have an impact on the forecasting NDVI models associated with crop response? (3) Can NDVI temporal series trends be predicted with prediction models? (4) How long in advance can we forecast the crop NDVI associated with the changes in environmental conditions?

The main contribution of this paper is to analyze predictions associated with different univariate methods (smoothed and ARIMAX) and their combinations, comparing their behavior. The literature shows that a combination of the predictions made by different procedures improves the precision of the individual predictions. The idea of combining predictions implicitly assumes that it is not possible to identify the underlying process in a series, and that each prediction method is capable of capturing different aspects of the information available for prediction; hence, a combination of predictions made using different techniques is the most accurate approach.

The paper is organized as follows: Section 2 describes the estimation and forecast methodologies used and the satellite datasets. The results and discussion are shown in Section 3, and Section 4 describes the main conclusions of this work and further future research.

## 2. Materials and Methodology Applied to the Forecast and Estimation Model

The target of the present study was to analyze vegetation indices (NDVI) obtained by a MODIS satellite and to construct models for forecasting and explanatory purposes. We considered time-dependent models to enrich the information given by the satellites in a given time period for a seasonal evolution of the vegetation indices for two types of Mediterranean crops (maize and olive tree) of the agricultural producing Jarama region (Madrid, Spain) during the February 2013–December 2017 period (Figure 1).

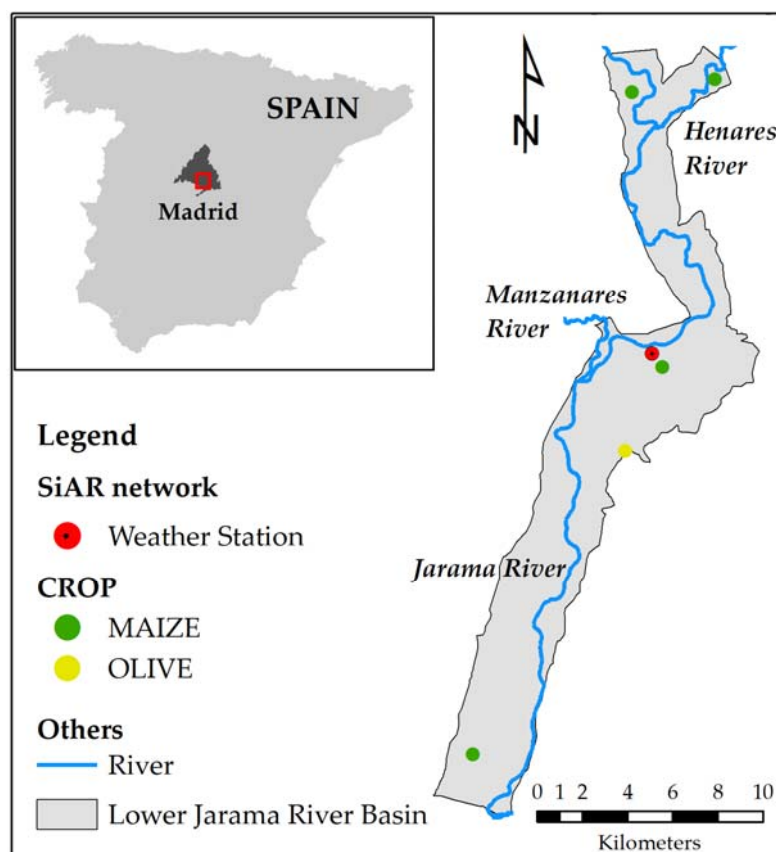


Figure 1. Lower Jarama River Basin, maize, olive and weather station location.

NDVI data were provided from a MOD13Q1 V006 at 250 m spatial resolution product derived from the MODIS sensor, produced at 16 day interval composite time series of the mean, which provide consistent spatial and temporal comparisons of vegetation status. The NDVI product chose the best available pixel value from all the acquisitions from the 16 day period, without clouds, low view angle and the highest NDVI value [33]. To eliminate possible errors in the NDVI data, the algorithm computes from surface reflectance corrected for molecular scattering, ozone absorption and aerosols, thus minimizing background variations and maintaining sensitivity in dense vegetation conditions [34].

The MOD13Q1 V006 NDVI products are available from the EarthExplorer website (<http://earthexplorer.usgs.gov/> (accessed on 20 November 2020)) provided by U.S. Geological Survey (USGS). The downloaded data were in HDF-EOS data format and sinusoidal projection. Therefore, they were reprojected using MODIS Reprojection Tool (MRT). To extract the NDVI data set, we used a vectorial mask for each plot (maize and olive), using a geographic information system (GIS). The GIS provides the ability to manage and analyze spatial and geographic data, thus improving the acquisition of data. Finally, we calculated the average value of the NDVI in each mask of the plots, as this is more representative of the state of the crops.

For the analysis of the data, four plots of maize and one of olive trees were selected, with areas larger than 7 hectares to ensure the consistency of the spatial resolution of the NDVI, which was 250m, and to eliminate errors due to the edge effect of the crop plots. In the case of maize, four different parcels were considered to carry out the study because in the Jarama region, the seeding and harvesting seasons of the maize crop can vary from one to two months, depending on the variety of maize and according to the use for which it is intended (human food or animal feed); the availability of water for irrigation; or the agroclimatic conditions of each year, such as rainfall and higher or lower temperatures.

In the database built for this paper, we added data monthly, from February 2013 to December 2017, with a total of 70 records for each plot. We chose the M02 (Arganda del Rey — UTM30N X457693; Y4462410) SiAR Network weather station for the agroclimatic data, because it is representative of the weather in the Lower Jarama River Basin and is near one of the maize plots [35] (Figure 1). The Agroclimatic Information System for Irrigation (SiAR) is an infrastructure of weather stations that captures, records and disseminates agroclimatic data. This information can be used to calculate the water demand of irrigation areas, which allows useful, rigorous and high-quality information that contributes to better planning, management, handling and control of irrigated farms to be obtained. The M02 station provides data from temperature ( $^{\circ}\text{C}$ ), wind speed (km/h), radiation ( $\text{W}/\text{m}^2$ ) and rainfall (mm) mainly.

### 2.1. B-Splines: Data Interpolation

The frequency in data sets obtained by satellites measures the appearance of missing observations due to climatological circumstances. Unobservable measures can be derived from cloudy atmospheric conditions when the satellite captures images on a given day or in a certain area. Unfortunately, time series methods for estimating purposes require full data bases, with no missing observations. Interpolation techniques allow us to construct new data points within the range of a discrete set of known data points. The simplest methods include linear interpolation, i.e., considering the two nearest available points and the linear expression between them to estimate missing observations, and polynomial interpolation, which is a generalization of the previous method considering a higher degree polynomial data. These methods have the advantage of being computationally fast, but the error is proportional to the square of the distance between the data points in the case of linear interpolation, and polynomial interpolations are computationally expensive and may exhibit oscillatory artifacts, especially at the end points (Runge's phenomenon) [36]. B-splines is generally used in this case, like global regression interpolation considering polynomial interpolation locally. The advantage of this method is the availability to generate maximum and minimum values, which is not possible when obtained by weighted

averages. These methods also provide smoother interpolants. The best results for the present study were obtained considering cubic B-splines when compared to Lagrange and Newton polynomials, avoiding Runge's phenomenon.

## 2.2. Time Series Analysis: Estimation Models

A useful methodology to study the evolution of time-dependent measures considering a period is time series analysis. This econometric model allows us to combine structural and time dimensions in an estimating model for forecast purposes. The main objective in time series analysis is to capture unobservable heterogeneity, between agents under study as well as in time. Several models are considered, attending to the structural patterns of the time series derived from the different crops in the study.

The simplest case can be modelled using the single exponential smoothing method, which is appropriate for series that move randomly above and below a constant mean with no trend and no seasonal patterns, and the error terms associated with the estimation model are uncorrelated. For the different crops considered in the study, we applied a double smoothing fit, appropriated for time series with a linear trend. Equation (2) presents the general mathematical model assumed for the variable in the study as a dynamic explanatory response in time. Let  $y_t$  be the observation of time series  $Y$  at period  $t \in \{1, \dots, T\}$ , then the single smoothed series  $S_t$  and the double smoothed series  $D_t$  are given by:

$$\begin{aligned} S_t &= \alpha y_t + (1 - \alpha)S_{t-1} \\ D_t &= \alpha S_t + (1 - \alpha)D_{t-1} \end{aligned} \quad (2)$$

where  $\alpha$  is the damping factor ( $0 < \alpha \leq 1$ ). Equation (3) presents the forecast mathematical model for a new observation at time period  $k$ .

$$\begin{aligned} y_{T+k} &= \left(2 + \frac{\alpha k}{1-\alpha}\right) S_T - \left(1 + \frac{\alpha k}{1-\alpha}\right) D_T \\ &= (2S_T - D_T + \frac{\alpha}{1-\alpha}(S_T - D_T)k) \end{aligned} \quad (3)$$

where  $\alpha$  is the weight used in the level component of the smoothed estimate.  $\alpha$  is like a moving average of the observations. The weights adjust the level of smoothing by defining the reaction of each component to current conditions. Lower weights offer less weight to recent data, which produces a smoother line. Higher weights give more weight to recent data, which produces a less smooth line.

More complex time-dependent structures can be modelled by dynamic regression methods and the composed autoregressive and moving averages method (ARIMA). These methods have been used widely in different application areas due to their flexibility and fitness ability for time-dependent data. ARIMA models were proposed by Box and Jenkins (1970) for stationary time series with linear autodependence [37]. They were developed with the aim of estimating and forecasting the behavior of the time data based on the previous performance considering autocorrelation in the error terms associated with the estimation model. These models were improved to include regressive variables known as the autoregressive integrated moving average with explanatory variables (ARIMAX). Equation (4) presents the general ARIMAX model, where  $B$  is the lag operator,  $d$  is the number of differences that must be applied to the series to make it stationary,  $\phi_i$  is the parameters of the  $p$  autoregressive part of the model,  $\theta_i$  is the parameters of the moving average part,  $x_j$  represents the model's exogenous variables,  $\lambda_j$  is the parameters of the exogenous part and  $\varepsilon_t$  is the error terms [37,38]. The error terms  $\varepsilon_t$  are generally assumed to be independent, identically distributed variables sampled from a normal distribution with zero mean.

$$(1 - B)^d y_t = a + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{h=1}^q \theta_h \varepsilon_{t-h} + \sum_{j=1}^r \lambda_j x_j + \varepsilon_t \quad (4)$$



In general, these models improve the prediction quality of the general model and allow us to analyze the impact of the explicative variables in the time series. In the present study, weather variables (precipitation, temperature and radiation) were considered as possible explicative measures in the estimation models. To evaluate the goodness of fit of the different proposed models, we considered the  $R^2$  determination coefficient and Akaike Criteria (AIC). To validate the proposed estimation models for forecasting purposes, we considered the Box Ljung Test, where the null hypothesis,  $H_0$ , is that data are independently distributed (i.e., the correlation in the population from which the sample is taken is 0, so that any observed correlation in the data results from randomness of the sampling process), and the alternate hypothesis,  $H_1$ , is that the model does show a lack of fit (i.e., data are not independently distributed and show serial correlation). When the p-values do not detect serial correlation at the chosen lags, this test does not reject the null hypothesis; thus, the time series are not autocorrelated [38,39].

### 2.3. Forecast Averaging: Prediction Model

Different forecast methods, from intuitive and subjective to more complex quantitative models, aim to develop accurate predictions. Furthermore, considering that the forecast related to a given magnitude can be developed by different agents and derived from various methods, the consideration of a prediction composed by the combination of several predictions is relevant. Since procedures differ in their philosophy, computational cost, complexity and accuracy, the selection of alternative methods or their combination results in a difficult task, the use of evaluation measures based on forecast errors being usual in practice. The main idea of the model is to use the concept of predictive combination, which has proven to be an effective method in the predictive literature. Several methods can be used to forecast the same variable, and the literature shows that a combination of predictions made through different processes can improve the precision of a single prediction. The idea of a combination of predictions implicitly assumes that the model cannot identify a number of basic processes, and that each prediction method can capture different aspects of information that can be used for prediction, so predictions made with different technologies are the most accurate forecasts.

Therefore, a combination of predictions derived from different methods usually improves the forecast accuracy. In this study, we combined predictions obtained by the proposed methods by means of forecast averaging. This method has the advantage of combining the relevant information of the time series given the different elements captured by each estimation model. The resulting prediction is expressed as a weighted average of the predictions resulting from the selected relevant estimation models, enriching the final forecast [40–42]. Among different approaches described in the literature, we selected the weighted average forecast procedure:

- Least Squares Weights (LS), computed as a regression of predicted values to real values and employing regression coefficients as weights for the prediction combinations [43].
- Mean Squared Error (MSE), assigning higher weight values to predictions associated with the model with smaller average quadratic error. MSE weighting, proposed by Stock-Watson (2001), compares the individual forecasts with the actual values over a forecast period [44]. The MSE (Equation (5)) of each forecast is computed and used to form individual forecast weights (Equation (6)):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5)$$

$$w_i = \frac{\frac{1}{MSE_i^k}}{\sum_{j=1}^N \frac{1}{MSE_j^k}} \quad (6)$$

$k$  is used to raise the MSE to different power;  $k = 1$  is the most used power, and yields a weight based on the ratio of each forecast's MSE to the total of all the MSEs.

To forecast procedures, we applied an inside proof-rolling sample, and in order to evaluate the proposed methods, we applied the MSE as a forecast sample measure.

#### 2.4. Econometric Data: Weather Variables, Outliers and Satellite Measures

Influence causes or factors in a process require adequate research to avoid non-desirable effects and control their dissemination. The cause search allows us to observe changes and study their mechanisms to generate new hypotheses and design plans to modify or mitigate their effects. When constructing an explanatory model, relationships between causes and a given effect can be represented in such a way that the proportion of each cause to the effect can be computed. The aim of this paper was to find a mathematical model that relates vegetation indices for different types of crops with possible explanatory variables. Additionally, it is important to detect the presence of outliers in the given effect. Their presence can be derived from multiple causes: a single time instant in which an outstanding value is obtained; a period in which the observed behavior of the time series changes drastically or returns gradually to the previous level. To consider all the possible factors that can provoke effects in the vegetation indices obtained by satellite, we considered as part of the explanatory model the following components:

- Weather variables, namely, temperature (T), precipitation (P), humidity (H) and radiation (R), were measured in the same time period as the satellite measures;
- Autoregressive measures obtained through the lags of the time series in the studio;
- Intervention measures to deal with outliers presented in the time series.

On the other hand, satellite measures obtained by MODIS are collected fortnightly. To apply time series modelling techniques, data must be aggregated in monthly frequencies.

We considered the following aggregation models for the five types of crops (Figure 2): time series obtained by the average values of available data in each month (Figure 2a); time series obtained by the maximum values of available data in each month (Figure 2b); and time series obtained by the minimum values of available data in each month (Figure 2c).

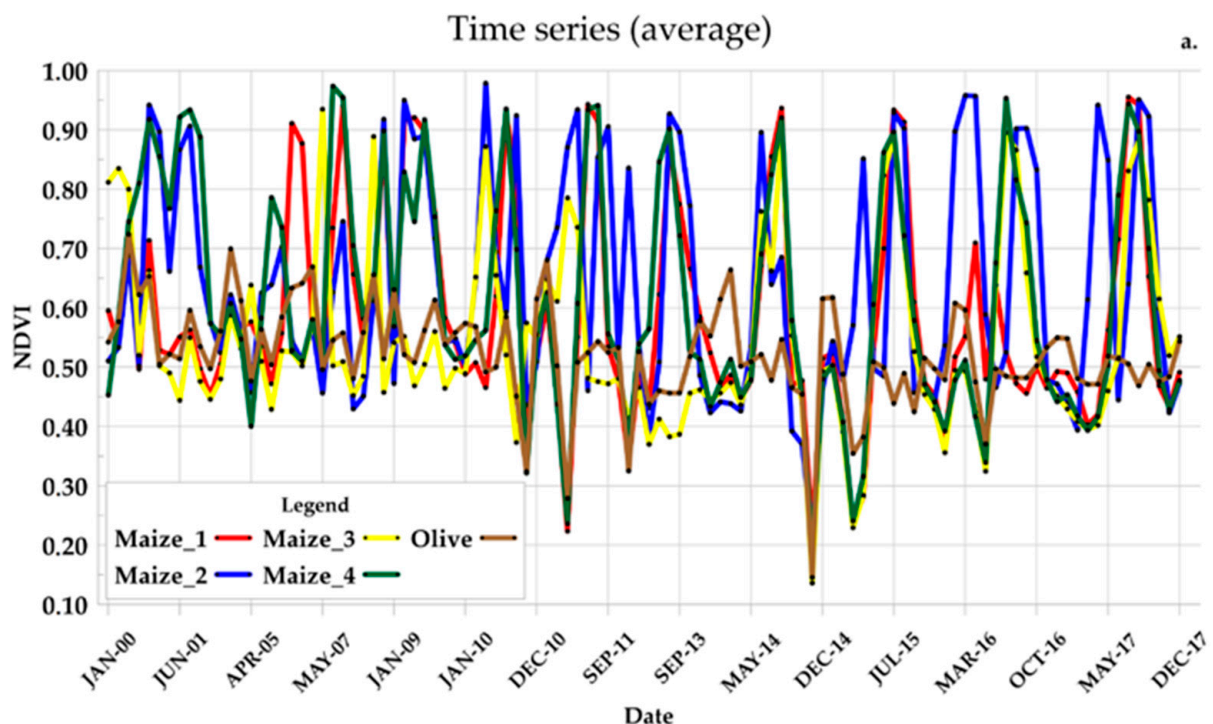
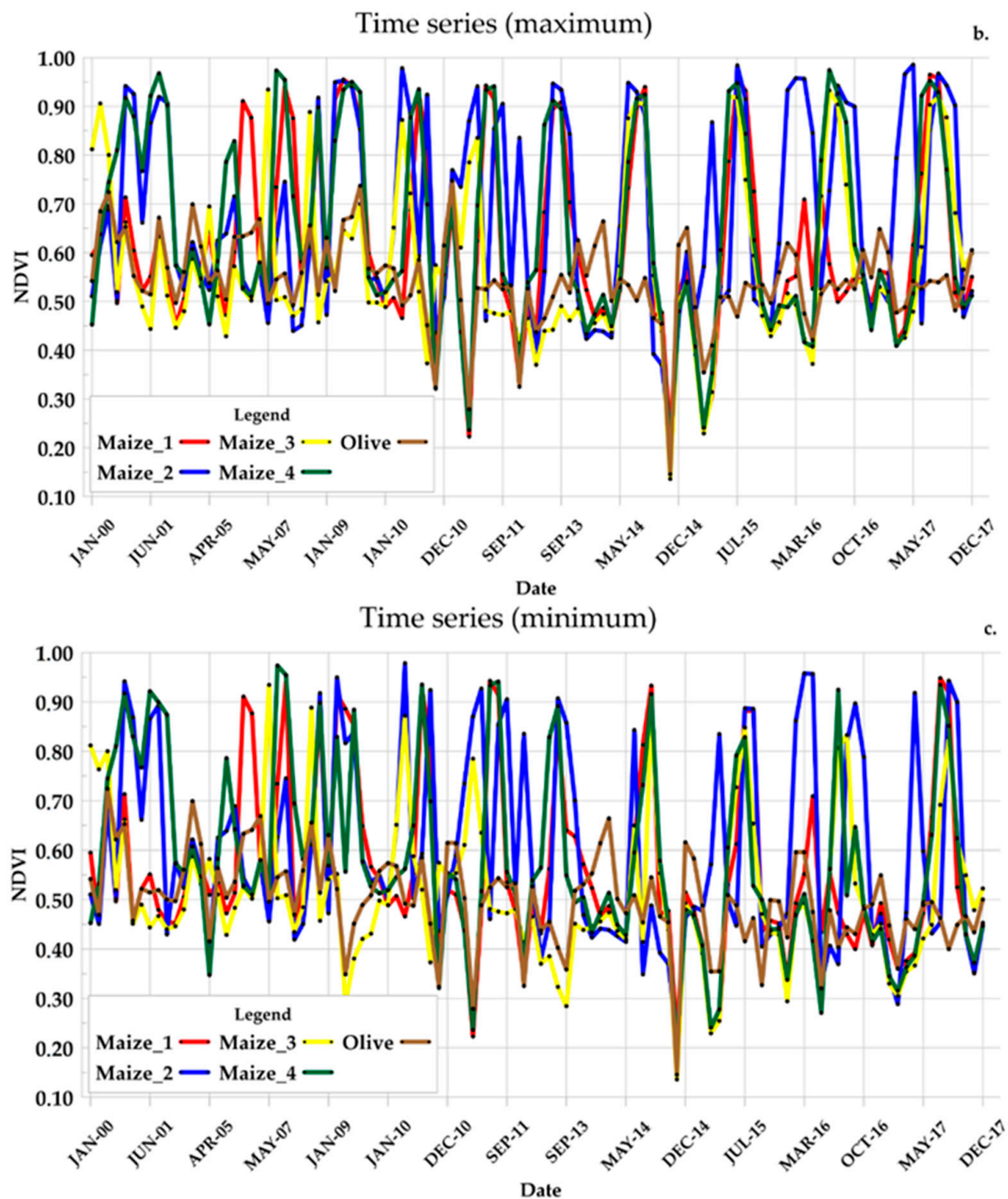


Figure 2. Cont.



**Figure 2.** Time series obtained by the average values (a), maximum values (b) and minimum values (c) of available data in each month.

In this paper, we consider the case obtained by the average values of the data available in each month, as there are no observable significant differences between the three proposed aggregation methods. In the case of unbalanced data, other aggregation measures can be used as an alternative, such as the median value. In our case, data were balanced, and the variation between the minimum and maximum values was small; therefore, the average was able to appropriately represent the sample population.



### 3. Results and Discussion

Using time series analysis, we developed forecasting and estimation models for the behavior of NVDI using satellite measures. We considered the time series obtained by the average values of available data in each month. However, the same methodology can be applied for the series obtained considering the maximum and minimum values. We considered an aggregated final model to forecast the time series values in the short term, considering 10 months of horizon planning (forecast values from January 2018 to September 2018). For model validation, we used the cross-validation method; divided the samples into training time series, comprising the available data from January 2013 to January 2015; and tested time series from February 2015 to December 2017. In consequence, we present final model equations avoiding non-significant parameters. The motivation for using cross-validation techniques is that when we fit a model, we fit it to the training data set. Without cross validation, we will only obtain information about how the model works with the sample data. Ideally, we want to see the performance of the model when new data are available based on the accuracy of the prediction.

Exponential smoothing models and ARIMAX were combined to include the meteorological measurement variables, namely, temperature (T), precipitation (P), humidity (H) and radiation (R), whose prediction errors were greater than the combination of their predictions using different methods.

For the elaboration of forecasts of the different crops with the ARIMAX structure, considering the dynamic structure of the time series and by means of transference functions, the variables that have an impact on their behavior must be identified, and in turn, the methodology used for the construction of the prediction model. The first step is the development of the iterative process to analyze the variables whose coefficients are significant to explain the behavior of the response variable.

In all the ARIMAX models that we present below, the meteorological measurements are included, and after adjusting and verifying the significance of the coefficients, the results of the diagnosis are presented below.

#### 3.1. Type 1 Maize

Double smoothing and ARIMAX models were applied to estimate the type 1 maize. Weather measures were found to be significantly influential in the models, contributing to the explanatory power of the proposed models. Equation (7) presents the estimation model for type 1 maize considering the interpolation obtained by means of the single exponential smoothing method.

$$\begin{aligned} S_t &= 0.87 y_t + (1 - 0.87)S_{t-1} \\ D_t &= 0.87 S_t + (1 - 0.87)D_{t-1} \end{aligned} \quad (7)$$

The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (6) is 0.709, and the Root Mean Squared Error (RMSE) is 0.111. The associated mean value for the model presented in Equation (7) is 0.5284, and the trend is  $-0.0014$ . Since  $\alpha = 0.87$  is the weight used in the level component of the smoothed estimate, the high weight gives more weight to the recent data.

Considering the dynamic structure of the time series and by means of transference functions, Equation (8) presents the estimation model for type 1 maize. The best selected model was ARIMAX(2,0,0)X(1,0,0) considering the smallest AIC value.

$$Y_t = 0.39 + 0.0116T + 0.117P + 0.657Y_{t-1} + 0.356Y_{t-2} + 0.414Y_{t-12} + \varepsilon_t \quad (8)$$

where  $Y_t$  is type 1 maize at time  $t$ ,  $T$  is the average temperature and  $P$  is the precipitation. All estimated parameters were significant at the 0.05 level. The SSR associated value for the model presented in Equation (8) is 0.629, and the Akaike info criterion (AIC) is  $-1.369$ . We can observe the positive impact of temperature and precipitation over type 1 maize and a dynamic dependence. Ljung Box autocorrelation proof validates the hypothesis of white noise in the residuals of the model, confirming the adequacy and reliability of the estimation

model. Figure 3 (Equation (8)) presents the estimation model as a dynamic explanatory response in time and the residuals of the estimation model given by Equation (8).



Figure 3. Representation for Maize 1 variety, estimation model and residuals given by Equations (8)–(11).

A similar model considering the dynamic structure of the time series and by means of transference functions is presented in Equation (9). The selected model was  $ARI-MAX(2,0,0) \times (1,0,0)$  considering the smallest AIC value.

$$Y_t = 0.34 + 0.011R + 0.723Y_{t-1} + 0.351Y_{t-2} + 0.449Y_{t-12} + \varepsilon_t \quad (9)$$

where  $Y_t$  is type 1 maize at time  $t$  and  $R$  is radiation. All estimated parameters were significant at the 0.05 level. The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (9) is 0.608, and the Akaike info criterion (AIC) is  $-1.38$ . We can observe the positive impact of the radiation over type 1 maize and a dynamic dependence. Ljung Box autocorrelation proof validates the hypothesis of white noise in the residuals of the model, confirming the adequacy and reliability of the estimation model. Figure 3 (Equation (9)) presents the estimation model as a dynamic explanatory response in time and the residuals of the estimation model given by Equation (9).

A similar model considering the dynamic structure of the time series and by means of transference functions is presented in Equation (10). The selected model was  $ARI-MAX(2,0,0) \times (1,0,0)$ .

$$Y_t = 0.82 + 0.004H + 0.761Y_{t-1} + 0.372Y_{t-2} + 0.466Y_{t-12} + \varepsilon_t \quad (10)$$

where  $Y_t$  is type 1 maize at time  $t$  and  $H$  is humidity. All estimated parameters were significant at the 0.05 level. The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (10) is 0.598, and the Akaike info criterion (AIC) is  $-1.39$ . We can observe the positive impact of the average humidity over type 1 maize and a dynamic

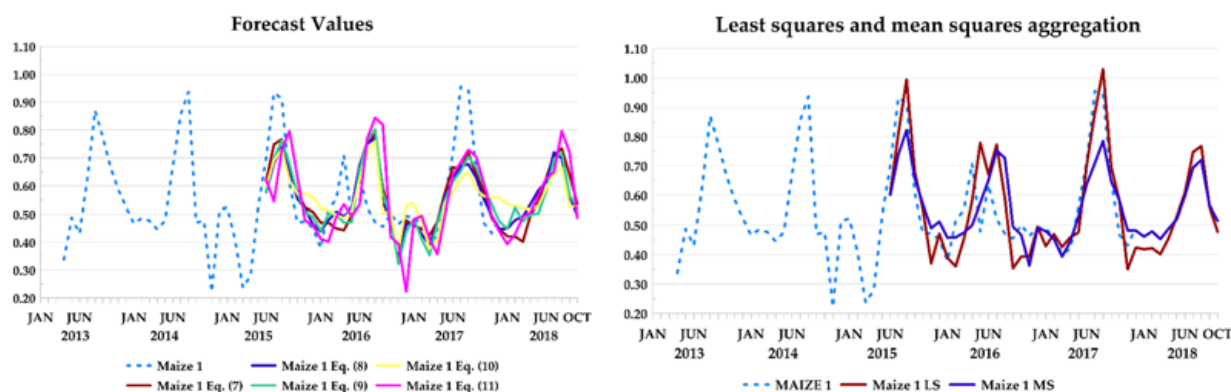
dependence. Ljung Box autocorrelation proof validates the hypothesis of white noise in the residuals of the model, confirming the adequacy and reliability of the estimation model. Figure 3 (Equation (10)) presents the estimation model as a dynamic explanatory response in time and the residuals of the estimation model given by Equation (10).

The last considered model considering the dynamic structure of the time series and by means of transference functions is presented in Equation (11). The selected model was ARIMAX(2,0,0)X(1,0,1).

$$Y_t = 0.55 + 0.708Y_{t-1} + 0.575Y_{t-2} + 0.547Y_{t-12} + \varepsilon_t + 0.533\varepsilon_{t-12} \quad (11)$$

where  $Y_t$  is type 1 maize at time  $t$ . All estimated parameters were significant at the 0.05 level. The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (11) is 0.563, and the Akaike info criterion (AIC) is  $-1.39$ . We can observe a dynamic dependence. Ljung Box autocorrelation proof validates the hypothesis of white noise in the residuals of the model, confirming the adequacy and reliability of the estimation model. Figure 3 (Equation (11)) presents the estimation model as a dynamic explanatory response in time and the residuals of the estimation model given by Equation (11).

The forecast model is constructed as a combination of models presented in Equations (7)–(11). Therefore, the dynamic structure of the type 1 maize time series is included as the effect of average temperature, precipitation, average humidity and radiation. The combination of the models was obtained through least squares and mean squares, with a resulting MSE of 0.01 and 0.011, respectively. Figure 4 presents the forecast values for a one-year time horizon given by the estimation models presented in Equations (7)–(11) and least squares and mean squares aggregation models. The associated MSEs for each estimation model presented in Equations (7)–(11) were 0.011, 0.017, 0.017, 0.019 and 0.019. The best forecast model attending to the best MSE value was obtained by least squares.



**Figure 4.** Forecast values for one-year time horizon given by the estimation models presented in Equation (7)–(11) and least squares and mean squares aggregation models.

### 3.2. Type 2 Maize

Double smoothing, ARIMA and ARIMAX models were applied to estimate the type 2 maize. Weather measures were found to be significantly influential in the models, contributing to the explanatory power of the proposed models. Equation (12) presents the estimation model for type 2 maize considering the interpolation obtained by means of the single exponential smoothing method.

$$S_t = 0.86 y_t + (1 - 0.86)S_{t-1} D_t = 0.86 S_t + (1 - 0.86)D_{t-1} \quad (12)$$

The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (12) is 1.418, and the Root Mean Squared Error (RMSE) is 0.157. The associated mean value for the model presented in Equation (12) is 0.647, and the trend is 0.001. The weight  $\alpha = 0.86$ , being high, offers more weight to recent data.

Considering the dynamic structure of the time series and by means of transference functions, Equation (13) presents the estimation model for type 2 maize. The best selected model was ARIMAX(2,0,0)X(0,0,0) considering the smallest AIC value.

$$Y_t = 0.45 + 0.099T + 0.606Y_{t-1} + 0.494Y_{t-2} + \varepsilon_t \quad (13)$$

where  $Y_t$  is type 2 maize at time  $t$  and  $T$  is the average temperature. All estimated parameters were significant at the 0.05 level. The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (12) is 1.343, and the Akaike info criterion (AIC) is  $-0.72$ . We can observe the positive impact of temperature over type 2 maize and a dynamic dependence. Ljung Box autocorrelation proof validates the hypothesis of white noise in the residuals of the model, confirming the adequacy and reliability of the estimation model. Figure 5 presents the estimation model as a dynamic explanatory response in time and the residuals of the estimation model given by Equation (13).

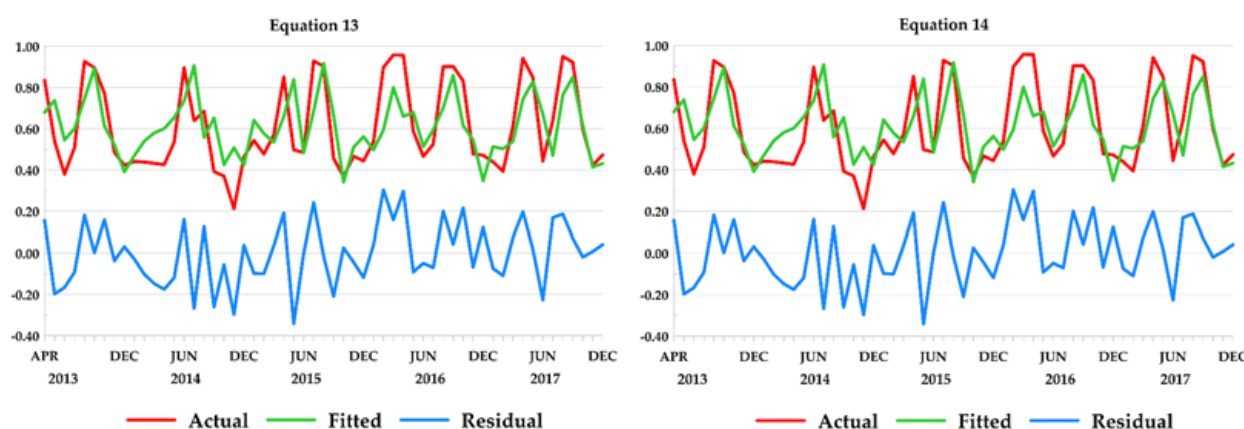


Figure 5. Representation for type 2 maize, estimation model and the residuals given by Equations (13) and (14).

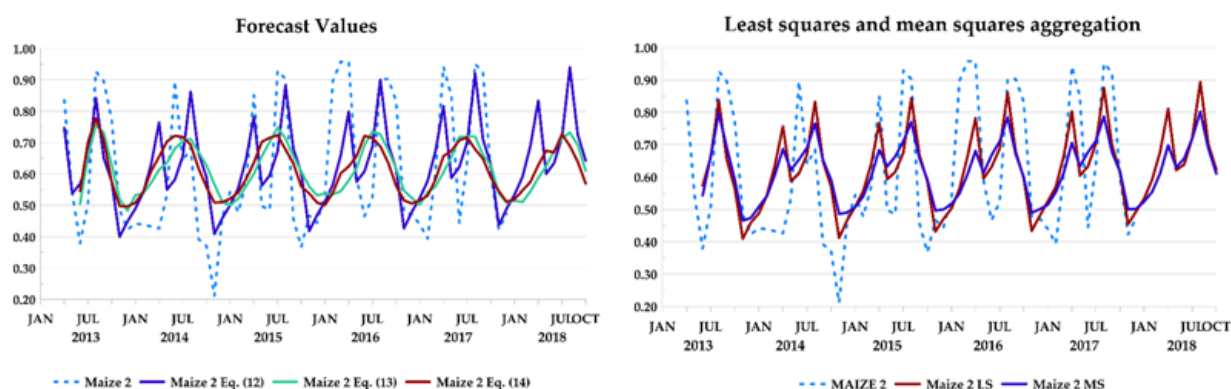
A similar model considering the dynamic structure of the time series and by means of transference functions is presented in Equation (14). The selected model was ARIMAX(2,0,0)X(0,0,0).

$$Y_t = 0.436 + 0.01R + 0.602Y_{t-1} + 0.499Y_{t-2} + \varepsilon_t \quad (14)$$

where  $Y_t$  is type 2 maize at time  $t$  and  $R$  is radiation. All estimated parameters were significant at the 0.05 level. The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (14) is 1.32, and the Akaike info criterion (AIC) is  $-0.73$ . We can observe the positive impact of the radiation over type 2 maize and a dynamic dependence. Ljung Box autocorrelation proof validates the hypothesis of white noise in the residuals of the model, confirming the adequacy and reliability of the estimation model. Figure 5 presents the estimation model as a dynamic explanatory response in time and the residuals of the estimation model given by Equation (14).

The forecast model is constructed as a combination of models presented in Equations (12)–(14). Therefore, the dynamic structure of the type 2 maize time series is included as the effect of average temperature and radiation. The combination of the models was obtained through least squares and mean squares, with a resulting MSE of 0.016 and 0.017, respectively. Figure 6 presents the forecast values for a one-year time horizon given by the estimation models presented in Equations (12)–(14) and least squares and mean squares aggregation models. The associated MSEs for each estimation model presented in Equations (12)–(14) were 0.026, 0.035 and 0.035. The best forecast model attending to the best MSE value was obtained by least squares.





**Figure 6.** Forecast values for one-year time horizon given by the estimation models presented in Equations (12)–(14) and least squares and mean squares aggregation models.

### 3.3. Type 3 Maize

Double smoothing, ARIMA and ARIMAX models were applied to estimate the type 3 maize. Weather measures were found to be significantly influential in the models, contributing to the explanatory power of the proposed models. Equation (15) presents the estimation model for type 3 maize considering the interpolation obtained by means of the single exponential smoothing method.

$$S_t = 0.75 y_t + (1 - 0.75)S_{t-1} D_t = 0.75 S_t + (1 - 0.75)D_{t-1} \quad (15)$$

The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (15) is 0.662, and the Root Mean Squared Error (RMSE) is 0.107. The associated mean value for the model presented in Equation (15) is 0.6, and the trend is 0.003. The weight  $\alpha = 0.75$ , being high, offers more weight to recent data.

Considering the dynamic structure of the time series and by means of transference functions, Equation (16) presents the estimation model for type 3 maize. The best selected model was ARIMAX(1,0,0)X(1,0,0) considering the smallest AIC value.

$$Y_t = 0.39 + 0.01T + 0.001P + 0.494Y_{t-1} + 0.185Y_{t-12} + \varepsilon_t \quad (16)$$

where  $Y_t$  is type 3 maize at time  $t$ ,  $T$  is the average temperature and  $P$  is the precipitation. All estimated parameters were significant at the 0.05 level. The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (16) is 0.688, and the Akaike info criterion (AIC) is  $-1.35$ . We can observe the positive impact of temperature and precipitation over type 3 maize and a dynamic dependence. Ljung Box autocorrelation proof validates the hypothesis of white noise in the residuals of the model, confirming the adequacy and reliability of the estimation model. Figure 7 presents the estimation model as a dynamic explanatory response in time and the residuals of the estimation model given by Equation (16).

A similar model considering the dynamic structure of the time series and by means of transference functions is presented in Equation (17). The selected model was ARIMAX(1,0,0)X(1,0,0).

$$Y_t = 0.813 + 0.005H + 0.502Y_{t-1} + 0.213Y_{t-12} + \varepsilon_t \quad (17)$$

where  $Y_t$  is type 3 maize at time  $t$  and  $H$  is the average humidity. All estimated parameters were significant at the 0.05 level. The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (17) is 0.71, and the Akaike info criterion (AIC) is  $-1.34$ . We can observe the positive impact of the average humidity over type 3 maize and a dynamic dependence. Ljung Box autocorrelation proof validates the hypothesis of white noise in the residuals of the model, confirming the adequacy and reliability of

the estimation model. Figure 7 presents the estimation model as a dynamic explanatory response in time and the residuals of the estimation model given by Equation (17).

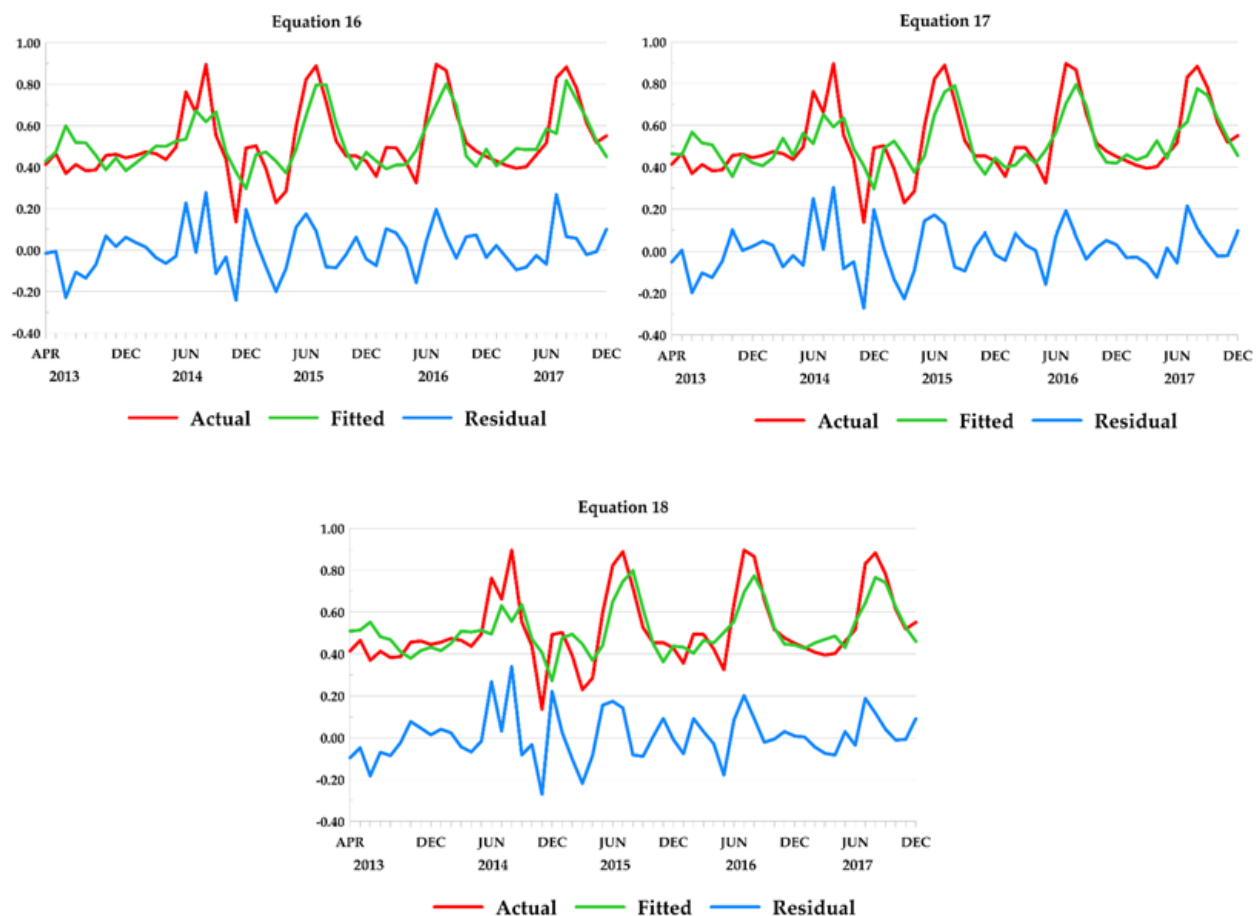


Figure 7. Representation for type 3 maize, estimation model and the residuals given by Equations (16)–(18).

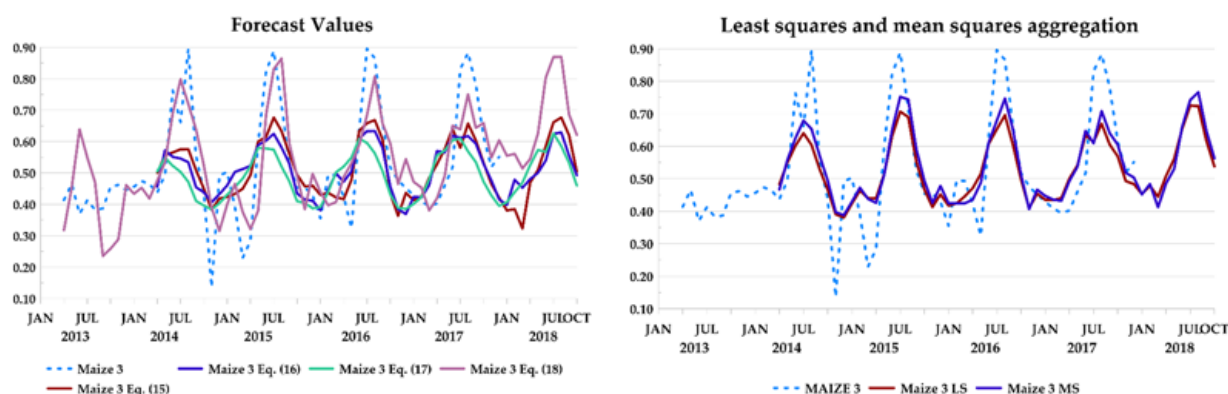
The last considered model taking into account the dynamic structure of the time series and by means of transference functions is presented in Equation (18). The selected model was ARIMAX(2,0,0)X(0,0,0).

$$Y_t = 0.33 + 0.01R + 0.529Y_{t-1} + 0.263Y_{t-2} + \varepsilon_t \quad (18)$$

where  $Y_t$  is type 3 maize at time  $t$  and  $R$  is radiation. All estimated parameters were significant at the 0.05 level. The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (18) is 0.736, and the Akaike info criterion (AIC) is  $-1.305$ . We can observe the positive impact of the radiation over type 3 maize and a dynamic dependence. Ljung Box autocorrelation proof validates the hypothesis of white noise in the residuals of the model, confirming the adequacy and reliability of the estimation model. Figure 7 presents the estimation model as a dynamic explanatory response in time and the residuals of the estimation model given by Equation (18).

The forecast model is constructed as a combination of models presented in Equations (15)–(18). Therefore, the dynamic structure of the type 3 maize time series is included as the effect of average temperature, precipitation, average humidity and radiation. The combination of the models was obtained through least squares and mean squares, with a resulting MSE of 0.009 and 0.011, respectively. Figure 8 presents the forecast values for a one-year time horizon given by the estimation models presented in Equations (15)–(18) and least squares and mean squares aggregation models. The associated MSEs for

each estimation model presented in Equations (15)–(18) were 0.012, 0.0182, 0.022 and 0.029. The best forecast model attending to the best MSE value was obtained by least squares.



**Figure 8.** Forecast values for one-year time horizon given by the estimation models presented in Equations (15)–(18) and least squares and mean squares aggregation models.

### 3.4. Type 4 Maize

Double smoothing, ARIMA and ARIMAX models were applied to estimate the type 4 maize. Weather measures were found significantly to be influential in the models, contributing to the explanatory power of the proposed models. Equation (19) presents the estimation model for type 4 maize considering the interpolation obtained by means of the single exponential smoothing method.

$$\begin{aligned} S_t &= 0.85 y_t + (1 - 0.85)S_{t-1} \\ D_t &= 0.85 S_t + (1 - 0.85)D_{t-1} \end{aligned} \quad (19)$$

The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (19) is 0.318, and the Root Mean Squared Error (RMSE) is 0.074. The associated mean value for the model presented in Equation (19) is 0.548, and the trend is  $-0.0005$ . The weight  $\alpha = 0.85$ , being high, offers more weight to recent data.

Considering the dynamic structure of the time series and by means of transference functions, Equation (20) presents the estimation model for type 4 maize. The best selected model was ARIMAX(2,0,1)X(0,0,0) considering the smallest AIC value.

$$Y_t = 0.025T + 0.001P + 0.003H + 0.703Y_{t-1} + 0.589Y_{t-2} + \varepsilon_t + 0.357\varepsilon_{t-1} \quad (20)$$

where  $Y_t$  is type 4 maize at time  $t$ ,  $T$  is the average temperature,  $P$  is the precipitation and  $H$  is the average humidity. All estimated parameters were significant at the 0.05 level. The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (20) is 0.401, and the Akaike info criterion (AIC) is  $-1.85$ . We can observe the positive impact of temperature, precipitation and humidity over type 4 maize and a dynamic dependence. Ljung Box autocorrelation proof validates the hypothesis of white noise in the residuals of the model, confirming the adequacy and reliability of the estimation model. Figure 9 presents the estimation model as a dynamic explanatory response in time and the residuals of the estimation model given by Equation (20).

A similar model considering the dynamic structure of the time series and by means of transference functions is presented in Equation (21). The selected model was ARIMAX(1,0,0)X(2,0,1).

$$Y_t = 0.028R + 0.426Y_{t-1} + 0.85Y_{t-12} + 0.174Y_{t-24} + 0.049\varepsilon_{t-12} + \varepsilon_t \quad (21)$$

where  $Y_t$  is type 4 maize at time  $t$  and  $R$  is radiation. All estimated parameters were significant at the 0.05 level. The Sum of Squared Residuals (SSR) associated value for the

model presented in Equation (21) is 0.298, and the Akaike info criterion (AIC) is  $-1.613$ . We can observe the positive impact of the radiation over type 4 maize and a dynamic dependence. Ljung Box autocorrelation proof validates the hypothesis of white noise in the residuals of the model, confirming the adequacy and reliability of the estimation model. Figure 9 presents the estimation model as a dynamic explanatory response in time and the residuals of the estimation model given by Equation (21).

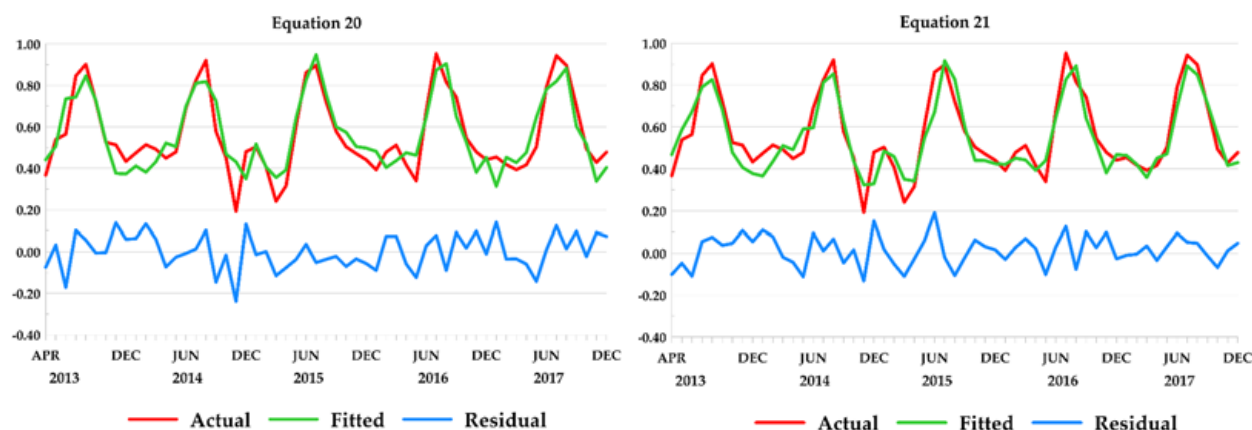


Figure 9. Representation for type 4 maize, estimation model and the residuals given by Equations (20) and (21).

The forecast model is constructed as a combination of models presented in Equations (19)–(21). Therefore, the dynamic structure of the type 4 maize time series is included as the effect of average temperature, precipitation, average humidity and radiation. The combination of the models was obtained through least squares and mean squares, with a resulting MSE of 0.001 and 0.002, respectively. Figure 10 presents the forecast values for a one-year time horizon given by the estimation models presented in Equations (19)–(21) and least squares and mean squares aggregation models. The associated MSEs for each estimation model presented in Equations (19)–(21) were 0.004, 0.008 and 0.008. The best forecast model attending to the best MSE value was obtained by least squares.

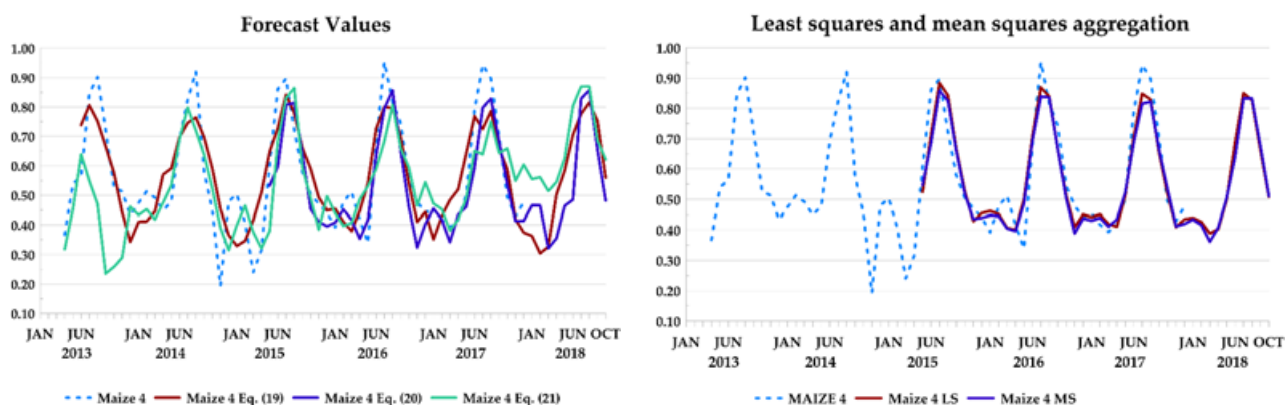


Figure 10. Forecast values for one-year time horizon given by the estimation models presented in Equations (19)–(21) and least squares and mean squares aggregation models.

### 3.5. Type of Maize Aggregate

Double smoothing and ARIMAX models were applied to estimate the type of maize aggregate. Weather measures were found to be significantly influential in the models, contributing to the explanatory power of the proposed models. Equation (22) presents the



estimation model for type of maize aggregate considering the interpolation obtained by means of the single exponential smoothing method.

$$\begin{aligned} S_t &= 0.85 y_t + (1 - 0.85)S_{t-1} \\ D_t &= 0.85 S_t + (1 - 0.85)D_{t-1} \end{aligned} \quad (22)$$

The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (22) is 1.4163, and the Root Mean Squared Error (RMSE) is 0.157. The associated mean value for the model presented in Equation (22) 0.49451, and the trend is 0.002152. The weight  $\alpha = 0.85$ , being high, offers more weight to recent data.

Equation (23) presents the estimation model for type of maize aggregate. The best selected model was ARIMAX(2,0,0)X(0,0,0) considering the smallest AIC value.

$$Y_t = 0.34 + 0.0122R + 0.67Y_{t-1} + 0.354Y_{t-2} + \varepsilon_t \quad (23)$$

where  $Y_t$  is type of maize aggregate at time  $t$ , and  $R$  is the radiation. All estimated parameters were significant at the 0.05 level. The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (23) is 0.496374, and the Akaike info criterion (AIC) is  $-1.720534$ . We can observe the positive impact of the radiation over type of maize aggregate and a dynamic dependence.

Ljung Box autocorrelation proof validates the hypothesis of white noise in the residuals of the model, confirming the adequacy and reliability of the estimation model. Figure 11 presents the estimation model as a dynamic explanatory response in time and the residuals of the estimation model given by Equation (23).

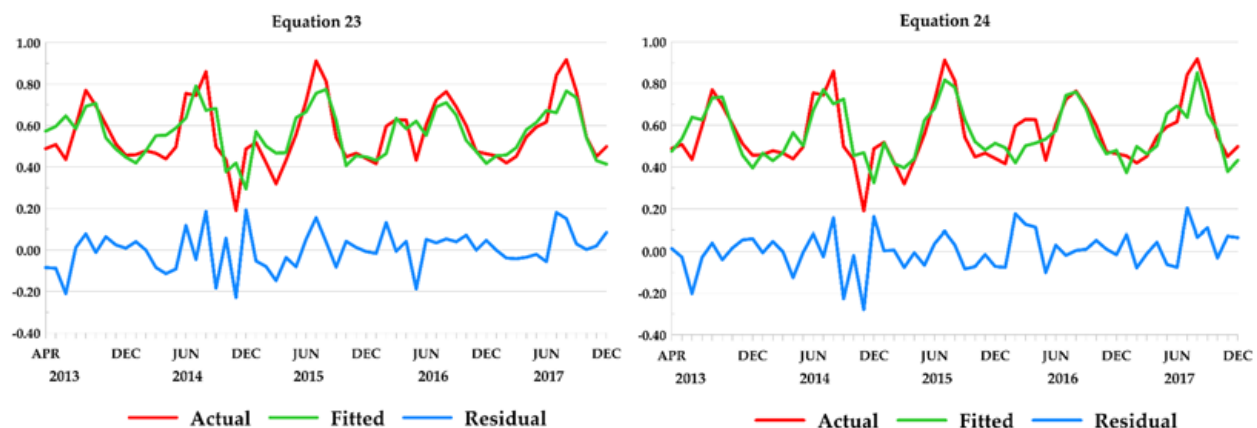


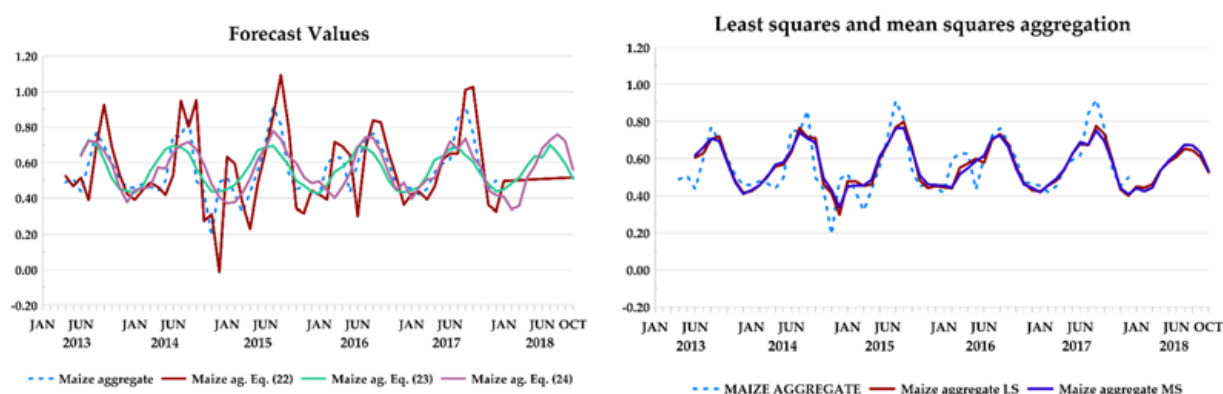
Figure 11. Representation for type of maize aggregate, estimation model and the residuals given by Equations (23) and (24).

A similar model considering the dynamic structure of the time series and by means of transference functions is presented in Equation (24). The selected model was ARIMAX(2,0,0)X(0,0,0).

$$Y_t = 0.022T + 0.0011 P + 0.041H + 0.473Y_{t-1} + 0.365Y_{t-2} + \varepsilon_t \quad (24)$$

where  $Y_t$  is type of maize aggregate at time  $t$  and  $R$  is radiation. All estimated parameters were significant at the 0.05 level. The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (24) is 0.47194, and the Akaike info criterion (AIC) is  $-1.73$ . We can observe the positive impact of the radiation over type of maize aggregate and a dynamic dependence. Ljung Box autocorrelation proof validates the hypothesis of white noise in the residuals of the model, confirming the adequacy and reliability of the estimation model. Figure 11 presents the estimation model as a dynamic explanatory response in time and the residuals of the estimation model given by Equation (24).

The forecast model is constructed as a combination of models presented in Equations (22)–(24). Therefore, the dynamic structure of the type of maize aggregate time series is included as the effect of average temperature, precipitation, average humidity and radiation. The combination of the models was obtained through least squares and mean squares, with a resulting MSE of 0.091 and 0.011, respectively. Figure 12 presents the forecast values for a one-year time horizon given by the estimation models presented in Equations (22)–(24) and least squares and mean squares aggregation models. The associated MSEs for each estimation model presented in Equations (22)–(24) were 0.026, 0.012 and 0.013. The best forecast model attending to the best MSE value was obtained by least squares.



**Figure 12.** Forecast values for one-year time horizon given by the estimation models presented in Equations (22)–(24) and least squares and mean squares aggregation models.

### 3.6. Olive

Double smoothing, ARIMA and ARIMAX models were applied to estimate olive crops. Weather measures were found to be significantly influential in the models, contributing to the explanatory power of the proposed models. Equation (25) presents the estimation model for olive considering the interpolation obtained by means of the single exponential smoothing method.

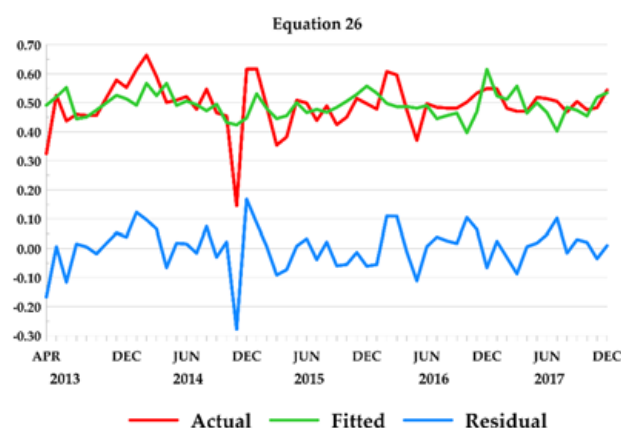
$$\begin{aligned} S_t &= 0.14 y_t + (1 - 0.14)S_{t-1} \\ D_t &= 0.14 S_t + (1 - 0.14)D_{t-1} \end{aligned} \quad (25)$$

The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (25) is 0.354, and the Root Mean Squared Error (RMSE) is 0.078. The associated mean value for the model presented in Equation (25) is 0.503, and the trend is 0.001. The weight  $\alpha = 0.14$ , being low, offers less weight to recent data.

Considering the dynamic structure of the time series and by means of transference functions, Equation (26) presents the estimation model for olive crops. The best selected model was ARIMAX(1,0,0)X(0,0,0) considering the smallest AIC value.

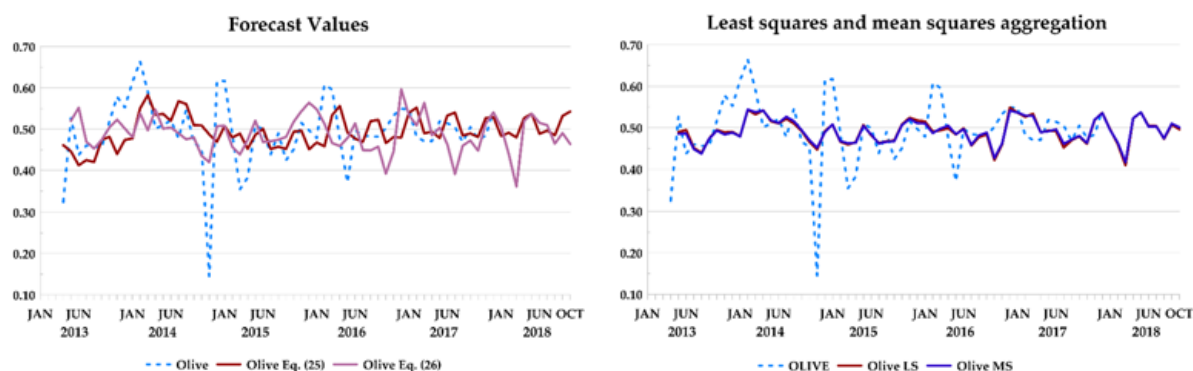
$$Y_t = 0.011P + 0.06H + 0.09R + 0.217Y_{t-1} + \varepsilon_t \quad (26)$$

where  $Y_t$  is olive crops at time  $t$ . All estimated parameters were significant at the 0.05 level. The Sum of Squared Residuals (SSR) associated value for the model presented in Equation (26) is 0.314, and the Akaike info criterion (AIC) is  $-2.18$ . We can observe a dynamic dependence. Ljung Box autocorrelation proof validates the hypothesis of white noise in the residuals of the model, confirming the adequacy and reliability of the estimation model. Figure 13 presents the estimation model as a dynamic explanatory response in time and the residuals of the estimation model given by Equation (26).



**Figure 13.** Representation for olive, estimation model and the residuals given by Equation (26).

The forecast model is constructed as a combination of models presented in Equations (25) and (26). The combination of the models was obtained through least squares and mean squares, with a resulting MSE of 0.003 and 0.003, respectively. Figure 14 presents the forecast values for a one-year time horizon given by the estimation models presented in Equations (25) and (26) and least squares and mean squares aggregation models. The associated MSEs for each estimation model presented in Equations (25) and (26) were 0.005 and 0.006. The best forecast model attending to the best MSE value was obtained by least squares.



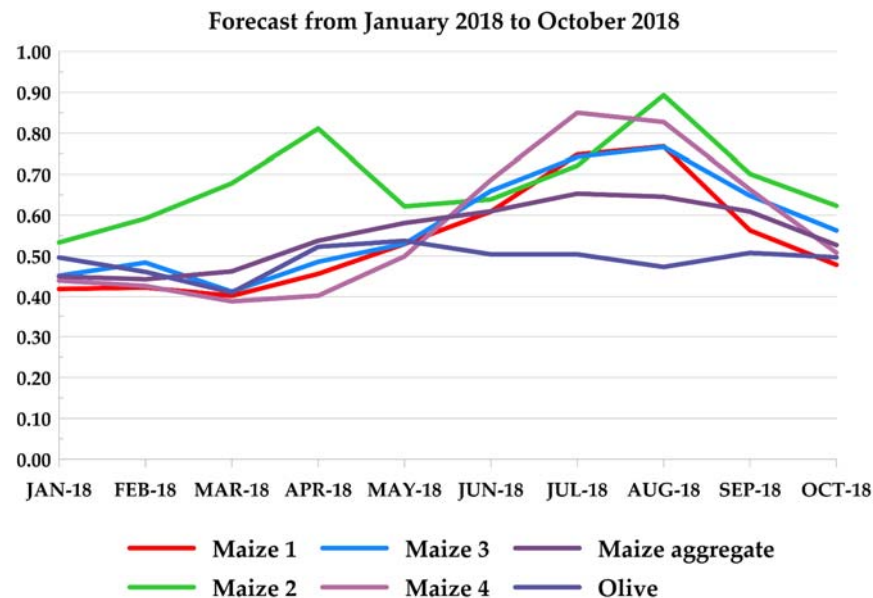
**Figure 14.** Forecast values for one-year time horizon given by the estimation models presented in Equations (25) and (26) and least squares and mean squares aggregation models.

The model equations present a positive contribution of the atmospheric variables T, P, H and R to the plant's growth. In the case of type 1 and 3 maize crops, the model considers lags in periods  $t = 1$ ,  $t = 2$  and  $t = 24$  because they have been maintained over several years, whereas for type 2 and 4 maize crops and the aggregated maize crops, the model considers solely lags in periods  $t = 1$  and  $t = 2$  because years of fallow have been interspersed.

The corresponding MSE is quite low for each model, which represents a good accuracy of the model fitting. The trends of all the considered models are near to zero, which implies very few variations in the considered time sample. For maize crops, the variation fluctuates between  $-0.0005$  (Type 1) and  $0.003$  (Type 1):  $0.002$  for the aggregated maize crop model and  $0.001$  for olive crop.

The weight obtained from double smoothing models ( $\alpha \in [0.75; 0.87]$ ) associated with maize crops is positive and near to one, which implies more weight to recent data; this is consistent with the phases of rapid evolution in the development of maize cultivation. In the case of olive trees, since it is a perennial tree crop with less physiological changes, the dependency of recent data is lower ( $\alpha = 0.14$ ).

As an example, we divided the sample into a training time series, comprising the available data from January 2013 to January 2015, and a tested (validation) time series from February 2015 to December 2017. Data from January 2018 to December 2018 were forecast out of the given sample. Figure 15 and Table 1 show the coefficients applied for the 2018 monthly prediction of the different types of crops.



**Figure 15.** Forecast out of sample for the first ten months of 2018 for four types of maize crops, olive and aggregated maize.

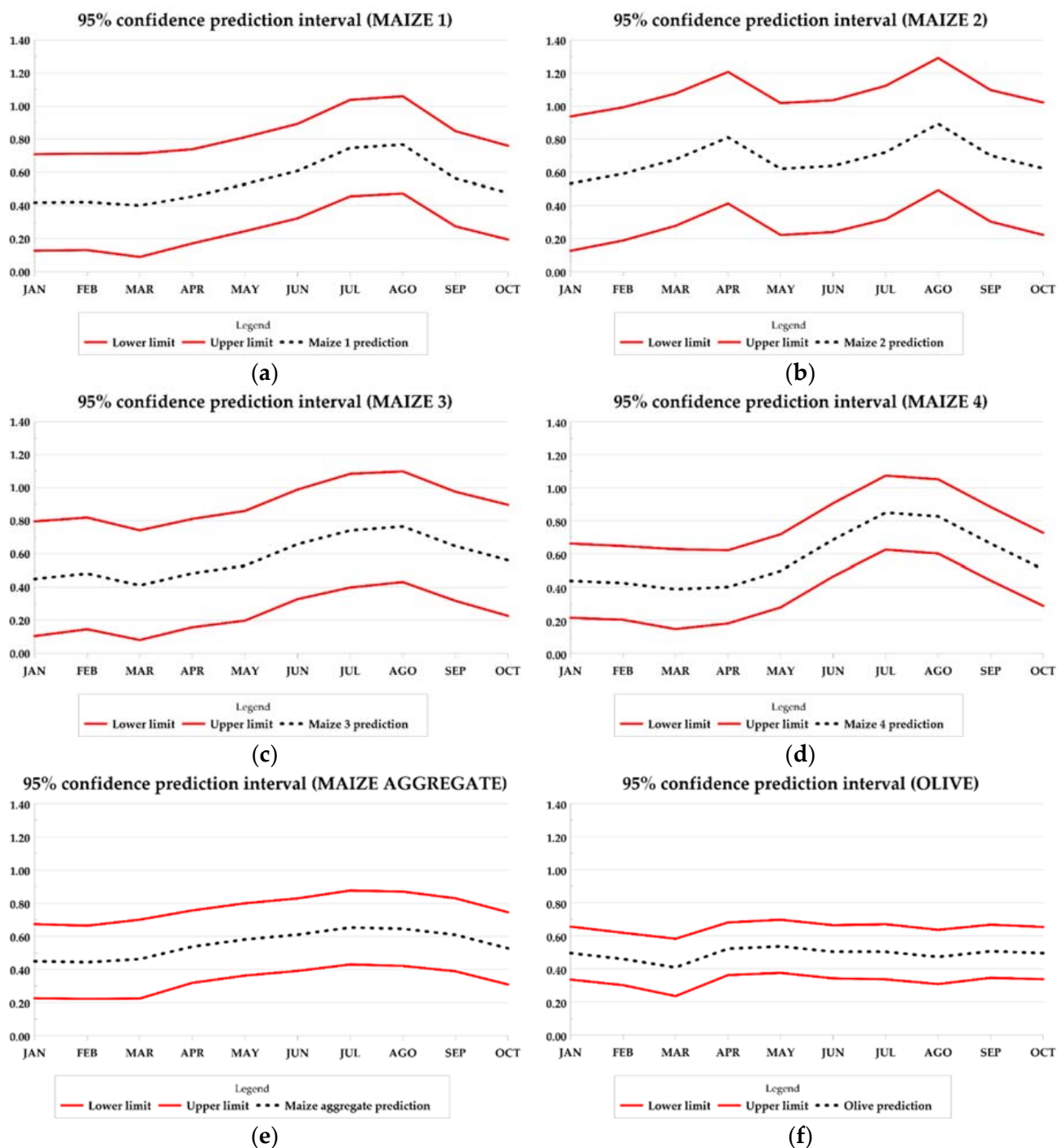
**Table 1.** Forecast for type 1–4 maize crops and olive crops from January 2018 to October 2018.

Time Period	Maize Type 1	Maize Type 2	Maize Type 3	Maize Type 4	Maize Aggregate	Olive
01/2018	0.418990	0.533135	0.451573	0.439748	0.449730	0.496241
02/2018	0.422720	0.591989	0.483806	0.426376	0.442843	0.461002
03/2018	0.402227	0.678137	0.412611	0.388669	0.462257	0.410330
04/2018	0.456171	0.811104	0.485727	0.402678	0.537400	0.522606
05/2018	0.529588	0.621669	0.530057	0.499155	0.580861	0.537193
06/2018	0.609008	0.638714	0.659564	0.686565	0.609717	0.504093
07/2018	0.747731	0.721217	0.742334	0.849985	0.652915	0.504024
08/2018	0.767648	0.893007	0.766180	0.827189	0.645175	0.473098
09/2018	0.562713	0.700946	0.647841	0.662895	0.608800	0.507662
10/2018	0.478441	0.623085	0.562679	0.508142	0.526868	0.496602

Finally, we display the associated forecast confidence intervals for  $\alpha = 0.05$  in Figure 16. These are the confidence intervals for the predictions shown in Table 1 for each of the crops: maize varieties, aggregated maize and olive.

In addition, it would be interesting to incorporate NDVI data obtained on the ground to check if the accuracy of these models can be improved. The NDVI derived from satellite sources and on the ground are not directly comparable; the NDVI data sets produced from these sources are frequently similar [45]. In this sense, the technological advances in unmanned vehicle systems (UAVS) that acquire NDVI data at low altitude [46] are of great help to monitor vegetation and crops with high spatial and radiometric precision, because they eliminate the effects of the atmosphere in the acquisition of reflectivity values in red and infrared.





**Figure 16.** Subfigures (a–d) presents maize crop types (1)–(4) forecast intervals with the associated  $\alpha = 0.05$ . Subfigure (e) presents aggregate maize crop types (1)–(4) forecast intervals with the associated  $\alpha = 0.05$ . Subfigure (f) presents olive crop type forecast intervals with the associated  $\alpha = 0.05$ .

#### 4. Conclusions

The results of this study presented above show a forecast model constructed as a combination of models applied to monitor crop dynamics, more specifically, a forecast algorithm applied to vegetation indices (NDVI) time series data closely related to the type and crop state derived from optical remote sensing.

The combination of predictions improves the precision of the individual predictions, and this work analyzed predictions associated with different univariate methods (smoothed and ARIMAX) and their combinations, comparing their behavior.

By comparing the different forecast models combinations, it was found that the equations of the models present a positive contribution of the climatological parameters of T, P, H and R for the growth of the plants. The considered climatological parameters are within the favorable ranges for the growth of the crops. Extreme climatic situations that could affect the normal development of the crops were not considered. This issue will be further developed in the future.

Following this, the equation models revealed that the temporal variable ( $Y_{t-1}$ ,  $Y_{t-2}$ ,  $Y_{t-12}$ ,  $Y_{t-24}$ ) included in each model depends on the type of crop, the interval being greater in the case of maize crops because they are seasonal crops and have a faster evolution. Furthermore, the weighting from the double smoothing model is high ( $\alpha = 0.85$ ). In the case of olive trees, as it is a perennial crop with fewer transformations, the dependence on the temporal variable is lower. Finally, considering the results obtained in this study, we highlight the following conclusions:

(1) Because each prediction method is capable of capturing different aspects of the information available for prediction, in this study, we offered a forecast algorithm combining different univariate methods (smoothed and ARIMAX) and their combinations for forecasting vegetation indices (NDVI) time series derived from optical remote sensing data.

(2) The accuracy of the prediction of the results, estimated as a weighted average of the predictions resulting from the selected relevant estimation models, was proved to be relevant in the final forecast.

(3) The considered climatological variables (temperature, air humidity, precipitation and solar radiation) have a positive and coherent contribution with the temporal evolution of the vegetation indexes but with different influence weights, according to the type of cultivation. Although in this work, we only explored data from two types of crops, four maize crops of different cycles, one aggregate for all maize crops and one plot of olive trees, in the future, it would be important to extend our methodology to other agricultural areas with other types of crops.

(4) The results of the prediction models were validated for a time series of one year and are very useful to be applied in the short term, in the management and planning of agricultural activities or availability of water resources for irrigation. In order to incorporate these forecasting models into climate models or hydrological models that require time series of tens of years, more data must be collected (NDVI and environmental conditions) to advance in longer forecasting models.

**Author Contributions:** All authors made a significant contribution to the final version of the paper. Conceptualization, F.C.-C., and A.E.S.; Methodology, F.C.-C., A.E.S., C.S., and D.M.-C.; Supervision, F.C.-C., A.E.S., C.S., and D.M.-C.; Processing remote sensing data: F.C.-C., and D.M.-C.; Weather variables data, D.M.-C. and F.C.-C.; Forecast models, A.E.S. and C.S.; Interpretation of results, F.C.-C., A.E.S., C.S., and D.M.-C.; Writing—original draft, F.C.-C.; Writing—review and editing, F.C.-C., A.E.S., C.S., and D.M.-C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Acknowledgments:** This study was financially supported by Agencia Estatal de Investigación (PID2019-108311GB-I00/AEI/10.13039/501100011033); PRODEBAT Spanish Ministry of Education and Science. Reference: 502 PID2019-106254RB-I00; Ministry of Economy and Competitiveness within the Retos-Collaboración 2014 program: “Smart-Hydro: Sistema Inteligente para optimizar el uso de agua en agricultura” (RTC-2014-2367-5); PDR-18-CAMEVAR: project, co-financed by the European Union through the European Agricultural Fund for Rural Development (EAFRD)—Europe invests in rural areas; MAPAMA and the Community of Madrid through IMIDRA (PDR-CM 2014-2020).

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

## References

- Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. Crop evapotranspiration: Guidelines for computing crop requirements. *Irrig Drain* **1998**, *300*, D05109. [\[CrossRef\]](#)
- Piao, S.; Cui, M.; Chen, A.; Wang, X.; Ciais, P.; Liu, J.; Tang, Y. Altitude and temperature dependence of change in the spring vegetation green-up date from 1982 to 2006 in the Qinghai-Xizang Plateau. *Agric. For. Meteorol.* **2011**, *151*, 1599–1608. [\[CrossRef\]](#)
- Ding, Y.; Xu, J.; Wang, X.; Peng, X.; Cai, H. Spatial and temporal effects of drought on Chinese vegetation under different coverage levels. *Sci. Total Environ.* **2020**, *716*, 137166. [\[CrossRef\]](#) [\[PubMed\]](#)
- Luo, N.; Mao, D.; Wen, B.; Liu, X. Climate change affected vegetation dynamics in the northern Xinjiang of China: Evaluation by SPEI and NDVI. *Land* **2020**, *9*, 90. [\[CrossRef\]](#)
- Zhang, X.; Ge, Q.; Zheng, J. Impacts and lags of global warming on vegetation in Beijing for the last 50 years based on remotely sensed data and phenological information. *Chin. J. Ecol.* **2005**, *2*, 123.
- Buitenwerf, R.; Rose, L.; Higgins, S.I. Three decades of multi-dimensional change in global leaf phenology. *Nat. Clim. Chang.* **2015**, *5*, 364–368. [\[CrossRef\]](#)
- Vicente-Serrano, S.M.; Gouveia, C.; Camarero, J.J.; Beguería, S.; Trigo, R.; López-Moreno, J.I.; Azorín-Molina, C.; Pasho, E.; Lorenzo-Lacruz, J.; Revuelto, J.; et al. Response of vegetation to drought time-scales across global land biomes. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 52–57. [\[CrossRef\]](#) [\[PubMed\]](#)
- Miah, M.G.; Abdullah, H.M.; Jeong, C. Exploring standardized precipitation evapotranspiration index for drought assessment in Bangladesh. *Environ. Monit. Assess.* **2017**, *189*, 1–16. [\[CrossRef\]](#) [\[PubMed\]](#)
- Brady, N.C.; Weil, R.R. *The Nature and Properties of Soils*; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2008.
- Beguería, S.; Vicente-Serrano, S.M.; Reig, F.; Latorre, B. Standardized precipitation evapotranspiration index (SPEI) revisited: Parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *Int. J. Climatol.* **2014**, *34*, 3001–3023. [\[CrossRef\]](#)
- García-Tejera, O.; López-Bernal, Á.; Orgaz, F.; Testi, L.; Villalobos, F.J. Analysing the combined effect of wetted area and irrigation volume on olive tree transpiration using a SPAC model with a multi-compartment soil solution. *Irrig. Sci.* **2017**, *35*, 409–423. [\[CrossRef\]](#)
- Ni, J.; Cheng, Y.; Wang, Q.; Ng, C.W.W.; Garg, A. Effects of vegetation on soil temperature and water content: Field monitoring and numerical modelling. *J. Hydrol.* **2019**, *571*, 494–502. [\[CrossRef\]](#)
- Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [\[CrossRef\]](#)
- Braun, M.; Herold, M. Mapping imperviousness using NDVI and linear spectral unmixing of ASTER data in the Cologne-Bonn region (Germany). In Proceedings of the Proc. SPIE 5239, Remote Sensing for Environmental Monitoring, GIS Applications, and Geology III, Barcelona, Spain, 13 February 2004; pp. 274–284.
- Quarmby, N.A.; Milnes, M.; Hindle, T.L.; Silleos, N. The use of multi-temporal NDVI measurements from AVHRR data for crop yield estimation and prediction. *Int. J. Remote Sens.* **1993**, *14*, 199–210. [\[CrossRef\]](#)
- Pettorelli, N.; Vik, J.O.; Mysterud, A.; Gaillard, J.M.; Tucker, C.J.; Stenseth, N.C. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends Ecol. Evol.* **2005**, *20*, 503–510. [\[CrossRef\]](#) [\[PubMed\]](#)
- Anderson, M.C.; Kustas, W.P.; Norman, J.M.; Hain, C.R.; Mecikalski, J.R.; Schultz, L.; González-Dugo, M.P.; Cammalleri, C.; d’Urso, G.; Pimstein, A.; et al. Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 223–239. [\[CrossRef\]](#)
- Subash, N.; Ram Mohan, H.S.; Banukumar, K. Comparing water-vegetative indices for rice (*Oryza sativa* L.)-wheat (*Triticum aestivum* L.) drought assessment. *Comput. Electron. Agric.* **2011**, *77*, 175–187. [\[CrossRef\]](#)
- Liu, X.; Zhu, X.; Li, S.; Liu, Y.; Pan, Y. Changes in Growing Season Vegetation and Their Associated Driving Forces in China during 2001–2012. *Remote Sens.* **2015**, *7*, 15517–15535. [\[CrossRef\]](#)
- Xue, J.; Su, B. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *J. Sensors* **2017**, *2017*, 1–17. [\[CrossRef\]](#)
- Karthikeyan, L.; Chawla, I.; Mishra, A.K. A review of remote sensing applications in agriculture for food security: Crop growth and yield, irrigation, and crop losses. *J. Hydrol.* **2020**, *586*, 124905. [\[CrossRef\]](#)
- Joiner, J.; Yoshida, Y.; Anderson, M.; Holmes, T.; Hain, C.; Reichle, R.; Koster, R.; Middleton, E.; Zeng, F.W. Global relationships among traditional reflectance vegetation indices (NDVI and NDII), evapotranspiration (ET), and soil moisture variability on weekly timescales. *Remote Sens. Environ.* **2018**, *219*, 339–352. [\[CrossRef\]](#)
- Liao, C.; Wang, J.; Pritchard, I.; Liu, J.; Shang, J. A spatio-temporal data fusion model for generating NDVI time series in heterogeneous regions. *Remote Sens.* **2017**, *9*, 1125. [\[CrossRef\]](#)
- Mutti, P.R.; Lúcio, P.S.; Dubreuil, V.; Bezerra, B.G. NDVI time series stochastic models for the forecast of vegetation dynamics over desertification hotspots. *Int. J. Remote Sens.* **2020**, *41*, 2759–2788. [\[CrossRef\]](#)
- Han, P.; Wang, P.X.; Zhang, S.Y.; Zhu, D.H. Drought forecasting based on the remote sensing data using ARIMA models. *Math. Comput. Model* **2010**, *51*, 1398–1403. [\[CrossRef\]](#)
- Jiang, B.; Liang, S.; Wang, J.; Xiao, Z. Modeling MODIS LAI time series using three statistical methods. *Remote Sens. Environ.* **2010**, *114*, 1432–1444. [\[CrossRef\]](#)

27. Fernández-Manso, A.; Quintano, C.; Fernández-Manso, O. Forecast of NDVI in coniferous areas using temporal ARIMA analysis and climatic data at a regional scale. *Int. J. Remote Sens.* **2011**, *32*, 1595–1617. [\[CrossRef\]](#)
28. Gonçalves, R.R.V.; Zullo, J.; Romani, L.A.S.; Nascimento, C.R.; Traina, A.J. Analysis of NDVI time series using cross-correlation and forecasting methods for monitoring sugarcane fields in Brazil. *Int. J. Remote Sens.* **2012**, *33*, 4653–4672. [\[CrossRef\]](#)
29. Rhee, J.; Im, J. Meteorological drought forecasting for ungauged areas based on machine learning: Using long-range climate forecast and remote sensing data. *Agric. For. Meteorol.* **2017**, *237–238*, 105–122. [\[CrossRef\]](#)
30. Reddy, D.S.; Prasad, P.R.C. Prediction of vegetation dynamics using NDVI time series data and LSTM. Model. *Earth Syst. Environ.* **2018**, *4*, 409–419. [\[CrossRef\]](#)
31. Ba, R.; Song, W.; Lovallo, M.; Lo, S.; Telesca, L. Analysis of Multifractal and Organization/Order Structure in Suomi-NPP VIIRS Normalized Difference Vegetation Index Series of Wildfire Affected and Unaffected Sites by Using the Multifractal Detrended Fluctuation Analysis and the Fisher–Shannon Analysis. *Entropy* **2020**, *22*, 415. [\[CrossRef\]](#)
32. Li, F.; Song, G.; Liujun, Z.; Yanan, Z.; Di, L. Urban vegetation phenology analysis using high spatio-temporal NDVI time series. *Urban For. Urban Green.* **2017**, *25*, 43–57. [\[CrossRef\]](#)
33. Didan, K.; Barreto Muñoz, A.; Solano, R.; Huete, A. *MODIS Vegetation Index User's Guide (MOD13 Series)*; Vegetation Index and Phenology Lab: Tucson, AZ, USA, 2015.
34. Choubin, B.; Soleimani, F.; Pirnia, A.; Sajedi-Hosseini, F.; Alilou, H.; Rahmati, O.; Melesse, A.M.; Singh, V.P.; Shahabi, H. Effects of drought on vegetative cover changes: Investigating spatiotemporal patterns. In *Extreme Hydrology and Climate Variability*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 213–222.
35. MAPAMA. SiAR: Sistema de Información Agroclimática para el Regadío. In Ministry of Agriculture, Fisheries and Food, Spain; Madrid. 2016. Available online: <http://www.siar.es> (accessed on 20 November 2020).
36. Fornberg, B.; Zuev, J. The Runge phenomenon and spatially variable shape parameters in RBF interpolation. *Comput. Math. Appl.* **2007**, *54*, 379–398. [\[CrossRef\]](#)
37. Box, G.; Jenkins, G. *Time Series Analysis: Forecasting and Control*; Holden-Day: San Francisco, CA, USA, 1970.
38. Hamilton, J. *Time Series Analysis*; Princeton University Press: Princeton, NJ, USA, 1994.
39. Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C. *Time Series Analysis: Forecasting and Control*, 4th ed.; San Francisco: Wiley, CA, USA, 2008.
40. Clemen, R.T. Combining forecasts: A review and annotated bibliography. *Int. J. Forecast* **1989**, *5*, 559–583. [\[CrossRef\]](#)
41. Makridakis, S.; Hibon, M. The M3-Competition: Results, conclusions and implications. *Int. J. Forecast* **2000**, *16*, 451–476. [\[CrossRef\]](#)
42. Timmermann, A. *Chapter 4 Forecast Combinations*; Elsevier: Amsterdam, The Netherlands, 2006; pp. 135–196.
43. Elliott, G.; Granger, C.; Timmermann, A. *Handbook of Economic Forecasting*; Elsevier: North Holland, The Netherlands, 2013; Volume 1.
44. Stock, J.H.; Watson, M.W. Vector Autoregressions. *J. Econ. Perspect.* **2001**, *15*, 101–115. [\[CrossRef\]](#)
45. Gozdowski, D.; Stępień, M.; Panek, E.; Varghese, J.; Bodecka, E.; Rozbicki, J.; Samborski, S. Comparison of winter wheat NDVI data derived from Landsat 8 and active optical sensor at field scale. *Remote Sens. Appl. Soc. Environ.* **2020**, *20*, 100409. [\[CrossRef\]](#)
46. Stow, D.; Nichol, C.J.; Wade, T.; Assmann, J.J.; Simpson, G.; Helfter, C. Illumination geometry and flying height influence surface reflectance and ndvi derived from multispectral UAS imagery. *Drones* **2019**, *3*, 55. [\[CrossRef\]](#)