Forecast of NDVI in coniferous areas using temporal ARIMA analysis and climatic data at a regional scale

Article *in* International Journal of Remote Sensing · March 2011

DOI: 10.1080/01431160903586765

CITATIONS
5

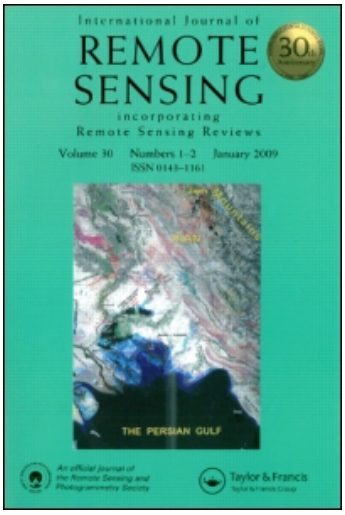
READS
271

3 authors:

Alfonso Fernandez-manso
Universidad de León
117 PUBLICATIONS **211** CITATIONS
SEE PROFILE

C. Quintano
Universidad de Valladolid
72 PUBLICATIONS **221** CITATIONS
SEE PROFILE

Oscar Fernández-Manso
Junta de Castilla y León
38 PUBLICATIONS **75** CITATIONS
SEE PROFILE



International Journal of Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713722504

Forecast of NDVI in coniferous areas using temporal ARIMA analysis and climatic data at a regional scale

A. Fernández-Manso[a]; C. Quintano[b]; O. Fernández-Manso[a]
[a] Agrarian Engineering Department, University of León, Av. Astorga s/n, Ponferrada, Spain [b] Electronic Technology Department, University of Valladolid, C/Francisco Mendizabal 1, Valladolid, Spain

Online publication date: 24 March 2011

To cite this Article Fernández-Manso, A. , Quintano, C. and Fernández-Manso, O.(2011) 'Forecast of NDVI in coniferous areas using temporal ARIMA analysis and climatic data at a regional scale', International Journal of Remote Sensing, 32: 6, 1595 — 1617
To link to this Article: DOI: 10.1080/01431160903586765
URL: http://dx.doi.org/10.1080/01431160903586765

International Journal of Remote Sensing
Vol. 32, No. 6, 20 March 2011, 1595–1617

Taylor & Francis
Taylor & Francis Group

Forecast of NDVI in coniferous areas using temporal ARIMA analysis and climatic data at a regional scale

A. FERNÁNDEZ-MANSO*†, C. QUINTANO‡ and O. FERNÁNDEZ-MANSO†
†Agrarian Engineering Department, University of León, Av. Astorga s/n, Ponferrada 24400, Spain
‡Electronic Technology Department, University of Valladolid, C/ Francisco Mendizabal 1, Valladolid 47014, Spain

(*Received 14 November 2005; in final form 23 November 2009*)

Important issues such as the prediction of drought, fire risk and forest disease are based on analysis of forest vegetation response. A method of forecasting the short-term response of forest vegetation on the basis of an autoregressive integrated moving average (ARIMA) analysis was designed in this study. We used 10-day maximum value composite (MVC) bands of the Normalized Difference Vegetation Index (NDVI) obtained from National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) data from 1993 to 1997. Using the theory of stochastic processes (Box–Jenkins), the MVC-NDVI series was analysed and a seasonal ARIMA (SARIMA) model was developed for forecasting NDVI in the following 10-day periods. The SARIMA model identified a moving-average regular term with a 10-day lag and an autoregressive 37 10-day period seasonal term with a one-season (1-year) component. The study also demonstrated a slight relationship between the NDVI and the precipitation level in some species of conifers by using climatic time series and the analysis of dynamic models and allowed us to elaborate an image of the immediate future NDVI for the study area (Castile and Leon, Spain).

1. Introduction

Remote sensing data have been produced for more than 30 years, providing a suite of information about ecological processes, and time series of different ecological variables have been obtained from this dataset. Forecasting deals with the generation of hypothetical realizations of a time series based on the observed values with the aim of obtaining accurate predictions of what could happen in the future. Different examples of forecasting based on remote sensing data exist. Luther *et al.* (1997) tested the potential of remote sensing to aid in insect defoliation forecasting and its associated impact on forest stands. Kumar (2004) used the pre-monsoon rainfall estimate from satellite data for predicting the monsoon onset over the Kerala coast. Nagler *et al.* (2008) applied a semi-distributed model to validate satellite snow cover data for short-term runoff forecasting. Relating agriculture variables, Sridhar *et al.* (1994) forecasted wheat production in India by using digital data from the Linear Imaging Self Scanner (LISS-I) onboard the Indian Remote Sensing Satellite (IRS-IB), and Almeida *et al.* (2006) studied a method of supporting sugarcane yield forecasting

*Corresponding author. Email: alfonso.manso@unileon.es

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160903586765

using vegetation spectral indices, principal component analysis and historic yield data in the state of São Paulo, Brazil.

Piwowar and Ledrew (2002) approached the study of time series for the detection of past environmental changes and the forecasting of future changes in environmental processes from two perspectives: (1) a descriptive and multivariate analysis that does not require the order of observation to be respected; and (2) a time series analysis, requiring that the values observed in time are studied with respect to their chronological order. Descriptive and multivariate methods have been widely used to monitor vegetation state. The simplest methods are based on comparing the values of an image with historic reference values (mean, typical deviation, maximums and minimums). Standardization in comparison with the maximum and the minimum of the period served to define the Vegetation Condition Index (VCI) (Kogan 1995, Liu and Kogan 1996, Unganai and Kogan 1998). This index, which is adequate for detecting anomalous trends in vegetation, has been widely used to determine fire-risk situations (Illera *et al.* 1997) or drought (Peters *et al.* 1991). Other descriptive approaches have obtained variables derived from the Normalized Difference Vegetation Index (NDVI) time series (annual integral, maximum–minimum difference, date of inflexion of the NDVI curve) and related them to productivity and seasonality (Paruelo and Lauenroth 1998).

Time series analysis is also a method commonly used to forecast future changes in environmental processes by using remote sensing data. Its two main goals are: (1) to identify the nature of the phenomenon represented by the sequence of observations; and (2) to predict future values of the time series variable (forecasting). Both of these objectives require the pattern of observed time series data to be identified and formally described. The autoregressive integrated moving average (ARIMA) analysis developed by Box and Jenkins (1976) allows just that; it has gained enormous popularity in many areas and research practice confirms its potential and flexibility (Bails and Peppers 1982). ARIMA models have been widely used in economics but there are few references to their use in remote sensing, despite their great potential for performing short-term predictions. Sridhar *et al.* (1994) forecasted wheat production in a predominantly unirrigated region of India by using a combination of forecasts from two different methodologies: the wheat yield–spectral relationship and time series analysis using ARIMA models. Piwowar and Ledrew (2002) used satellite images and autoregressive moving average (ARMA) models in their study of climatic change, demonstrating their usefulness in identifying changes and forecasting future situations with statistical confidence. Ji and Petters (2004) designed a new vegetation greenness forecast model capable of forecasting the vegetation status for cropland and grassland in advance, based on a seasonal ARIMA (SARIMA), and Riaño *et al.* (2007) demonstrated that SARIMA time series modelling is a powerful tool for forecasting potential burned areas and obtaining results with statistical significance.

Different models have also been developed to explain the relationship between climate and vegetation by using NDVI values. Farrar *et al.* (1994), Liu *et al.* (1994) and Richard and Poccard (1998) related the NDVI to climatic variables such as precipitation. Zhou *et al.* (2001), Gong and Ho (2003), Kaufmann *et al.* (2003) and Jarlan *et al.* (2005) studied NDVI–climate relationships from a regional and global scale, showing a clear dependency between these variables.

Our study evaluates the application of ARIMA models to NDVI time series as a potential tool for forecasting natural risks in forested areas of the Castile and Leon region (Spain), considering also the relationship between the NDVI and climate. Our

main purpose was twofold: (1) to confirm the usefulness of ARIMA NDVI time series modelling in understanding the nature of the NDVI time series and forecasting the future vegetation state in order to identify risk situations due to the water stress of vegetation; and (2) to understand better the climate–NDVI relationship and to investigate whether using climate variables such air temperature and precipitation level can improve the previously mentioned ARIMA NDVI time series model. We suggest that it is possible to forecast vegetation status by the statistical manipulation of present and historical records from the National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) NDVI and climate variables.

2. Materials

2.1 *Study area*

Castille and Leon (Spain) was the region chosen as the study area (figure 1). This region is located in the central northern area of Spain and covers 94 147 km^2. Its landscape is defined by a system of high plains (700–1100 m) in a large sedimentary

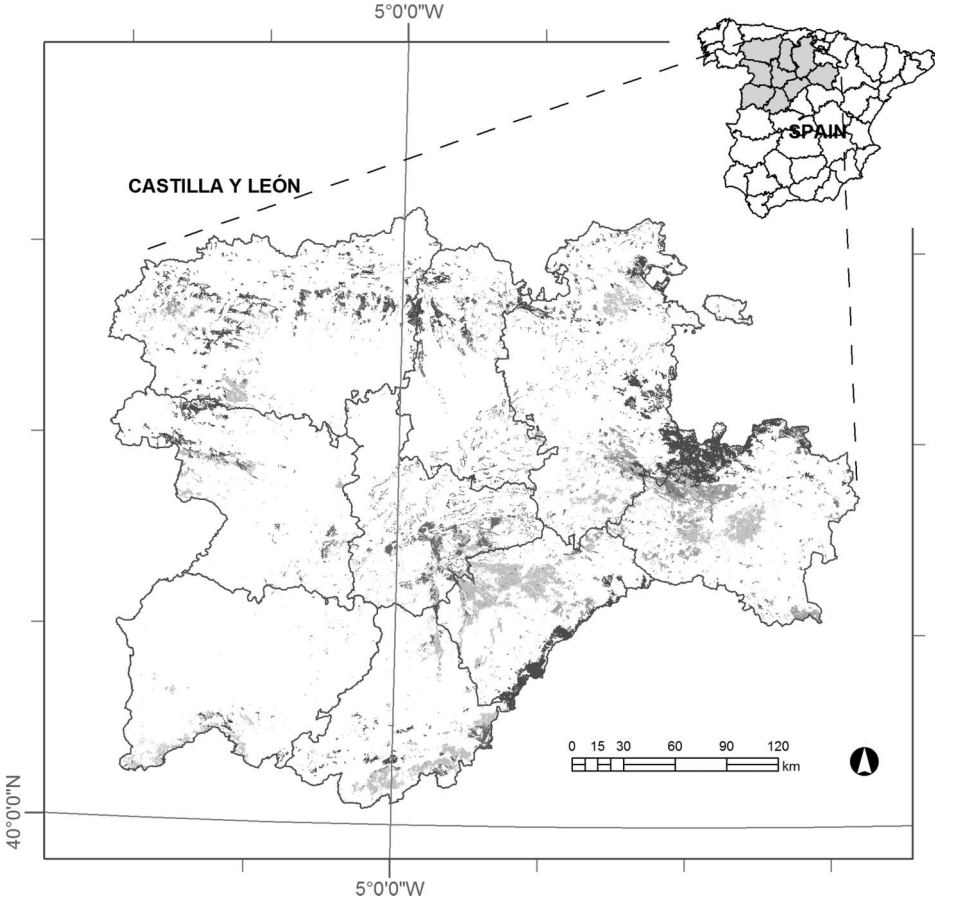


Figure 1. Study area and general conifer distribution.

basin whose limits are formed by different mountain ranges. This contrasted relief and its geographical situation mean a great diversity of ecology and landscape. A wide diversity in thermal situations can be found as well, though with an evident predominance of extreme temperatures with a mean of around 10–12°C in the central area. Precipitation within the region can be considered too low, as regards ecological needs; irregular, as regards annual quantity; and unequal, as regards spatial distribution. Precipitation quantity shows a gradient that increases with altitude and varies between 450 and 1600 mm.

The forestry land of Castille and Leon represents 50% of its total land area, whereas more than 50% of the forestry land is wooded land (ICONA 1990). The forests of pure conifers, incense juniper (*Juniperus thurifera* L.) and pines (*Pinus* spp.) cover 33% of the region. In this study six conifer species representative of the study area were chosen to construct the NDVI forecast models: *Pinus sylvestris* L., *Pinus nigra* L., *Pinus pinaster* Ait., *Pinus pinea* L., *Pinus halepensis* L. and *Juniperus thurifera* L.

2.2 *Dataset*

The work scale was 1: 2 000 000, which we considered sufficient for regional monitoring. Three types of information were used: NOAA-AVHRR images, vegetation information and climatic information. First, NOAA-AVHRR images on daily dates from 1 January 1993 to 31 December 1997 were used to obtain a 10-day period of maximum value composites of NDVI (MVC-NDVI). Second, the information from the Spanish Forestry Map (SFM) allowed us to elaborate a map of the types of vegetation at the work scale. The SFM, elaborated within the Spanish Second National Forest Inventory (NFI2) frame, stores the distribution of Spanish forestry masses on a 1:50 000 scale. Third, concerning climatic information, air temperature and precipitation level data were the climatic variables specifically considered. To define them, data from 240 air temperature stations and 740 precipitation level stations were used. These stations are part of the station network of the Spanish state-owned Meteorology Agency (AEMET, Agencia Estatal de METeorología). Changes in these climatic variables (from 1 January 1993 to 31 December 1997) were monitored every 10 days for similarity to MVC-NDVI.

3. Methods

3.1 *ARIMA analysis*

Time series are modelled mathematically as stochastic processes. Assuming that the observed value of a series at time t is a random extraction of an aleatoric variable defined at t, a series of data is considered as a sample of a vector of n aleatoric variables ordered in time $(z_1, \ldots, z_t, \ldots, z_n)$. The set of variables $\{z_t\}$, $t = 1, \ldots, n$, is called a stochastic process and the series observed is considered a realization of the process. The probabilistic structure of a stochastic process is based on the joint distribution of the n aleatoric variables $\{z_t\}$. The practical determination of the joint distribution of the process requires a large number of realizations to be observed. This quantification is greatly simplified when the joint distribution is a normal multivariate, as it will then be determined by the vector of means and the matrix of variances and co-variances between variables.

A general ARMA(p,q) model, the autoregressive process of moving averages of the order 'p, q', is expressed as follows:

$$(1 - \phi 1B - \ldots - \phi pBp)(1 - B)\mathrm{d}z_t = (1 - \theta 1B - \ldots - \theta qBq)a_t \quad (1)$$

or in compact notation:

$$\phi p(B)\nabla \mathrm{d}z_t = \theta q(B)a_t \quad (2)$$

where z_t are observations at time t; $\phi p(B)$ and $\theta q(B)$ are, respectively, autoregressive and moving average coefficients with B being the backward shift operator; p denotes the polynomial order of the autoregressive part of the model based on previous data; and q denotes the polynomial order of the moving average part of the model based on the difference from the average of the previous data. Finally, a_t is white noise or model residuals, assumed to follow a normal distribution with a mean zero and constant scale.

An ARIMA(p,q,d) model allows a series of observations to be described after they have been differentiated d times, to extract possible non-stationary sources. It is a matter of generalizing an ARMA(p,q) model, adding a differentiation of the order of d. In these models, seasonality can also appear, in which case the SARIMA model has to be considered. This model is denoted as SARIMA(p,q,d) $\times$ (P,Q,D)$_S$, where P and Q are, respectively, the orders of the seasonal autoregressive and moving average coefficients; D represents the deterministic seasonal trend to make the time series stationary and invariant; and S represents seasonality.

3.2 *Methodology stages*

The proposed methodology can be defined by five stages (figure 2).

Stage 1. Preparation of the dataset

This stage includes all the operations carried out as a prior stage to the proposed methodology; specifically, radiometric and geometric correction of the NOAA-AVHRR data, elaboration of a forest cover thematic map, elaboration of a series of climatic maps and determination of the NDVI for each conifer species considered.

(1) *Radiometric and geometric correction of the NOAA-AVHRR data*. These corrections were performed using the software developed by the Remote Sensing Laboratory of Valladolid University (LATUV). To obtain radiance values, the method proposed by Kaufman and Holben (1993) was used. Cloud-covered areas were identified (Delgado 1991) and an image with a spatial resolution of 1 $\times$ 1 km in Universal Transverse Mercator (UTM) coordinates was obtained. The geometric correction of original NOAA-AVHRR data, based on orbital models (Illera *et al.* 1995), showed a root mean square error (RMSE) of less than 1 pixel. NDVI values are influenced by the reflectance direction of the Earth's surface and the differences of the satellite's spatial resolution in each pixel. Moreover, it is necessary to consider the disturbances caused by the atmosphere in infrared bands, which tend to reduce the index. We used 10-day MVC to minimize these effects, although unfortunately they are not sufficient to eliminate the atmospheric disturbances. A 10-day period was selected because it is the closest to the cycle of the satellite.

(2) *Elaboration of a forest cover thematic map*. A raster map (UTM projection, 1 $\times$ 1 km) with the six conifer classes of interest (*Pinus sylvestris* L., *Pinus nigra* L., *Pinus pinaster* Ait., *Pinus pinea* L., *Pinus halepensis* L. and *Juniperus thurifera* L.) was obtained from the SFM.

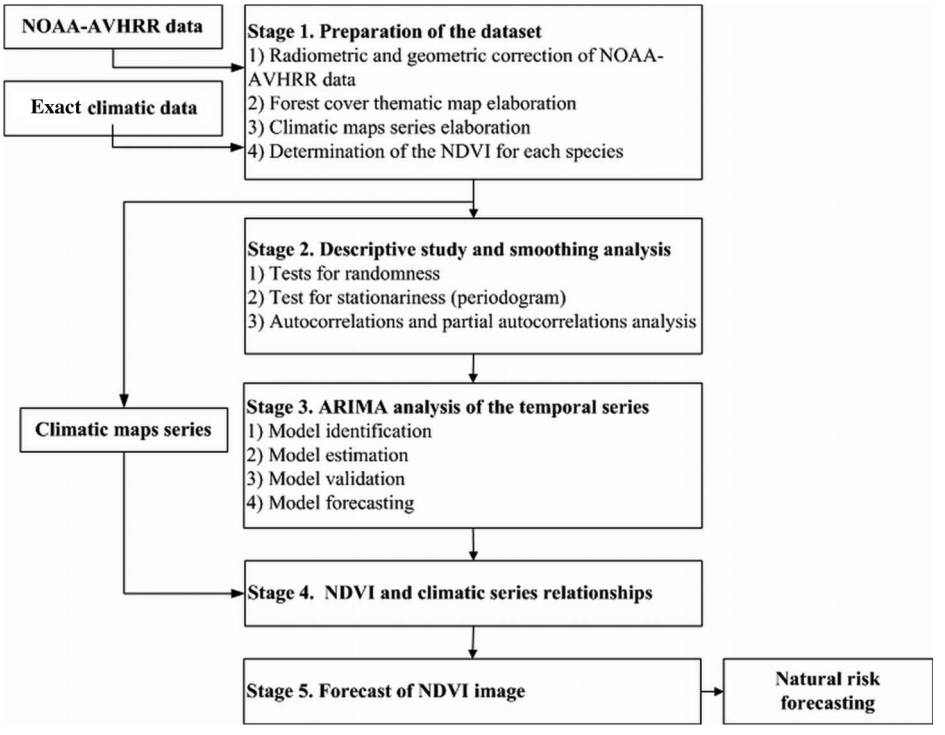


Figure 2. Flowchart of the methodology.

(3) *Elaboration of a series of climatic maps*. Interpolation of the climatic data used (air temperature and precipitation level) allowed us to obtain a raster map of these variables, matching the MVC-NDVI images. As regards the air temperature variable, a reciprocal distance squared interpolation (Chang 1973) of thermal values recorded at 240 AEMET stations was made to obtain maps of the temperatures, taking into account the relief effect using a Digital Terrain Model (DTM). The aim is for each value in the interpolated map to be representative of the air temperature experienced by the pixel (one square kilometre, the same resolution as NOAA-AVHRR data) during the 10-day period. The temperature variables, chosen for their phytoclimatic significance, were: mean (T_a), maximum (T_M) and minimum temperature (T_m), as well as the average of the maxima (T_{aM}) and the average of the minima (T_{am}). T_a was calculated as the 10-day mean of the daily average temperature; T_M and T_m as the 10-day maximum/minimum of the daily maximum/minimum. The interpolation was checked by relating the temperature values recorded at the AEMET meteorological stations to those obtained at the same point by the methodology. The maximum error was $\pm 1^{\circ}$C in 75% of the stations.

With regard to the precipitation variable, the rainfall fields of the whole territory were obtained from 740 AEMET stations by Thiessen tessellation (range 62.28–143.33 km^2, mean 127 km^2). Thiessen tessellation was chosen for its importance and extensive use. Aurenhammer (1991) describes Thiessen tessellation as one of the most fundamental data structures in computational geometry that are used in modelling natural

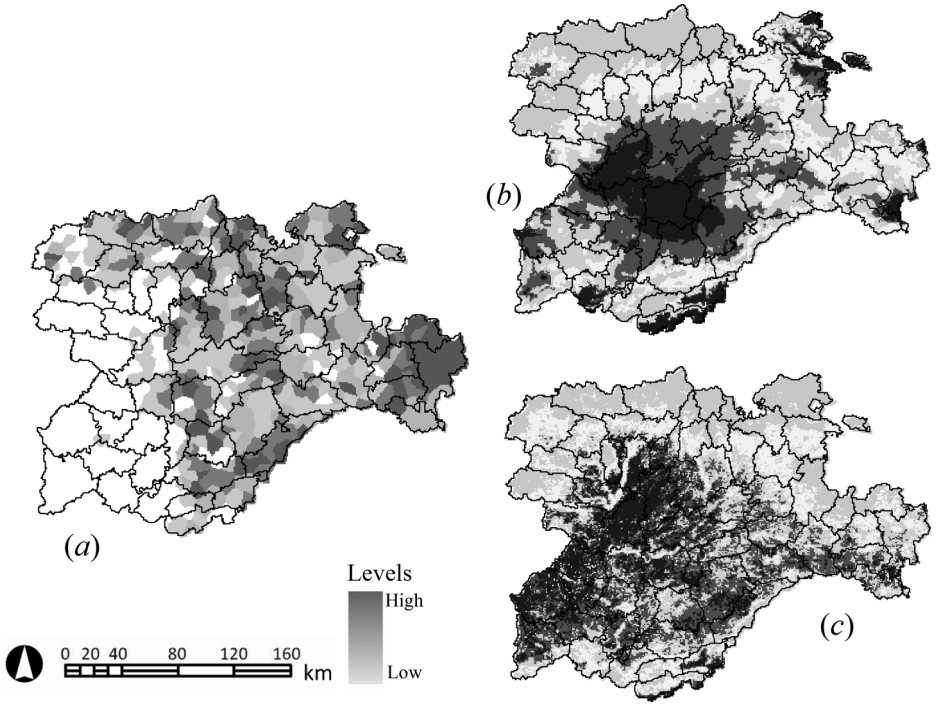


Figure 3. First 10-day period of June 1994: (*a*) precipitation levels; (*b*) average air temperature; (*c*) average maximum air temperature. The darker the colour, the higher the level.

phenomena, to investigate their mathematical, in particular, geometrical, combinatorial and stochastic properties and their computational representation. The precipitation variable was the 10-day accumulated precipitation level.

Finally, a series of 37 10-day compounded maps per year was obtained for each of those variables in a raster format (UTM projection, 1 km × 1 km). Figure 3 shows some examples of these maps.

(4) *Determination of the NDVI for each species*. From the forest cover thematic map obtained previously, the average time series corresponding to each of the six species of conifers considered was extracted from the MVC-NDVI historical image records. In this way, six series of 185 (37 × 5) observations were generated for the 1993–1997 period.

Stage 2. Descriptive study and smoothing of the series

A descriptive study, followed by smoothing, of each of the series obtained in the previous stage was performed as an analysis prior to modelling in order to check the randomness and stationarity of the data. To be able to construct a model of the time series the data must be non-random and stationary: if this second condition is not met, the appropriate transformations should be made (stage 3).

Randomness analysis guarantees that the series is not a random succession of numbers. Three statistical tests (a test for excessive runs above and below the median (RUNM); a test for excessive runs up and down (RUNS); and a Box–Pierre test) were

performed to test the null hypothesis of randomness of the six MVC-NDVI time series. However, a stationarity study also allows us to separate long-term variability from short-term variability by identifying the non-stationary structure. The identification of the non-stationary structure of the six MVC-NDVI time series was carried out in three steps: (1) determining whether it was necessary to transform the series to have constant variance; (2) determining the number of differences that must be applied to have a constant mean (generally 1 or 2); and (3) identifying whether seasonality existed and, if so, to proceed with the seasonal differentiation (Box *et al.* 1994).

Seasonality can be identified by using periodograms. A periodogram is the Fourier transform of the autocovariance function and allows the examination of frequency-domain models of an equispaced time series. From the point of view of data analysis, the type of structure in the autocovariance function indicates the location of peaks in the periodogram and these peaks indicate the dominant frequency for underlying cyclic models.

Stage 3. ARIMA analysis of the time series

The MVC-NDVI time series were analysed by using ARIMA models, taking the following steps: identification, estimation, validation and forecasting (Makridakis *et al.* 1998, Peña *et al.* 2001). The object of the first step, identification of the model, is to determine which model or models are best suited to the time series. Each model presents a characteristic autocorrelation function (ACF) and a distinctive partial autocorrelation function (PACF). The ACF is the linear correlation between the time series and the same time series with time lags. The PACF is the autocorrelation remaining after the ACF is accounted for. Specifically, the ACF and PACF from the observed data were compared with theoretical values for predetermined functions, with the closest matches preselected for further analysis in the next step.

In the second step, the estimation of the parameters of the ARIMA(p,q,d) model was carried out by selecting the most suitable model among the previously preselected ones. Furthermore, if the time series are not stationary (checked in stage 2), this estimation will be preceded by differentiation operations to transform the time series into a stationary one. To choose the definitive model, the following criteria were used: (1) the ACF and PACF functions; (2) a test of coefficient significance (with p-value below 0.05 it is statistically significant at a confidence level of 95%); (3) the errors made in the estimation of each model; and (4) five tests on the residuals. The magnitude of the errors was evaluated on the basis of the RMSE, mean absolute error (MAE) and mean absolute percentage error (MAPE). The best model is the one giving the lowest value. In addition, the slope was measured from the mean error (ME) and the mean percentage error (MPE): the best model had a value close to 0. The tests performed on the residuals were: RUNS; RUNM; a Box–Pierce test for excessive autocorrelation; a test for difference in the mean first half to the second half; and a test for the difference in variance first half to the second half.

Third, once the model was estimated, the results must be validated before using it for forecasting. For the fitted model to be acceptable, certain requisites that are implicitly or explicitly incorporated into its specification must be complied with. The purpose of the tests performed is to identify whether the residual series, which results from applying a model, is a series of random distribution data with a zero mean, which is equivalent to testing whether the model is capable of explaining practically all the variance, except for the term due to random ones (white noise

series). The following criteria were used in this study: (1) a study of the ACF and PACF of the residuals; (2) a study of the differences between the errors of the model in the estimation period and those of the validation period; and (3) a test for the suitability and randomness of the residuals. With reference to the first criterion, the residual series must be a white noise process with a zero mean. The ACF and PACF of the residuals must not have a determined form and no peak must exceed the bands of confidence, paying special attention to the first two peaks (other peaks can be allowed to exceed the limits). Regarding the second criterion, which was initially used to select the models, RMSE, MAE, MAPE, ME and MPE were calculated for the values of the model without the data to be validated and then with the validation data. If the results were considerably worse in the validation periods, this would mean that the model is not suitable for forecasting. Finally, considering the third criterion, six tests (RSME; RUNS; RUNM; a Box–Pierce test; a test for difference in mean first half to second half, a test for difference in variance first half to second half) were carried out to check the suitability of the residuals and three (RUNM, RUNS, Box–Pierce test) to study their randomness (with a p-value ≥ 0.10 it is not possible to reject the hypothesis that the residuals are random with a confidence level of 90%).

Lastly, after quantifying and validating the model, it became a useful instrument for short-term forecasts.

Stage 4. NDVI–climate relationships

To study the influence of the climatic variables considered on the inter-annual evolution of the NDVI, the six series of 185 observations of MVC-NDVI were compared to the climatic series of temperature and precipitation (185 observations, also) using dynamic models (a mixture of time series analysis and regression analysis).

Stage 5. Forecast of NDVI image

After stage 3, a valid model for short-term forecasting is ready to be used. Using this model and the MVC-NDVI images prior to a specific date, a new image including the estimated immediate future NDVI for the different species of conifers was obtained. This forecasted NDVI image was compared with the real one to verify the forecasting process.

4. Results

The NOAA-AVHRR dataset was adequately prepared and six MVC-NDVI time series (one for each species of conifer considered) were obtained. To determine their randomness (stage 2), three statistical tests were performed for each of them. The results showed (p-value < 0.01; confidence level 99%) that the time series studied were not random and therefore could be used to construct predictive models.

Regarding stationarity, the six time series of the conifers studied presented an initial non-stationary behaviour. As an example, figure 4(a) represents the ACF for *Pinus sylvestris* L. It showed a positive autoregressive (AR) structure with a very slow decrease. Analysing its periodogram (figure 4(b)), the cycles presenting a greater contribution to the variance could be identified. In the six time series considered, the occurrence was observed every 37 10-day periods, which determined a clear annual seasonality ($37 \times 10 = 370$), as expected. If the seasonality had been exactly periodic, it could have been eliminated from the time series as a deterministic

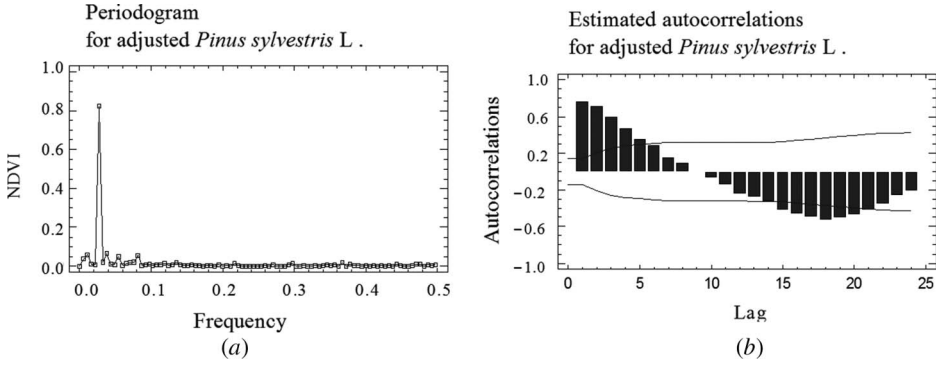


Figure 4. *Pinus sylvestris* L. MVC-NDVI time series: (*a*) ACF to test seasonality; (*b*) periodogram.

component. As it was only approximately constant, it had to be eliminated by taking the differences between the observations separated by the seasonal period (in all cases, a differentiation of order 1 had to be applied).

Four models elaborated from the combination of non-seasonal AR and non-seasonal moving-average (MA) parameters with seasonal AR (SAR) and seasonal MA (SMA) parameters were preselected for each MVC-NDVI time series. As described in §3.2, three criteria (the ACF and PACF functions, a test of coefficient significance and the errors made in the estimation of each model) were used to choose the most suitable model for each conifer species (stage 3). Figures 5(*a*)–5(*e*) and 5(*f*)–5(*j*) show, respectively, the ACF and PACF functions for each MVC-NDVI time series.

As an example, the four preselected models for *Pinus pinaster* L. were: (A) SARIMA(0,1,1) × $(0,1,1)_{37}$; (B) SARIMA(1,1,1) × $(0,1,1)_{37}$; (C) SARIMA(1,1,1) × $(1,1,1)_{37}$; and (D) SARIMA(0,1,1) × $(1,1,1)_{37}$. Table 1 shows the magnitude of the errors mentioned in §3.2 for each of these models. Model A was selected because it had the smallest RMSE, MAE and MAPE during the fitting phase. Table 2 summarizes the results of five tests run on the residuals to determine whether each model is adequate for the data. Note that model A passed four tests. As no tests are statistically significant at the 95% or higher confidence level, this model is probably adequate for the data. Considering all this information, model A [SARIMA(0,1,1) × $(0,1,1)_{37}$] was finally selected for *Pinus pinaster* L.

After analysing in a similar way the criteria for the rest of the conifer species considered, the results were mostly adjusted to a SARIMA(0,1,1) × $(0,1,1)_{37}$ with a moving average component in the seasonal and non-seasonal part. In *Pinus sylvestris* L. and *Pinus nigra* L., however, an AR component appeared in the seasonal part and the selected model was SARIMA(0,1,1) × $(1,1,1)_{37}$. Table 3 shows the results of the test of confidence significance for the selected models and table 4 shows the errors made in their estimations.

As indicated in the §3.2, three criteria were used to validate these selected models (the ACF and PACF of the residuals, the study of the differences between the errors of the model in the estimation period and those of the validation period, and tests for the suitability and randomness of the residuals). With reference to the first criterion, the ACF and PACF of the residuals for the six classes of conifer considered had no determined form and all the values were within the bands of confidence. Regarding the second criterion, the model was estimated from the first 175 data values and

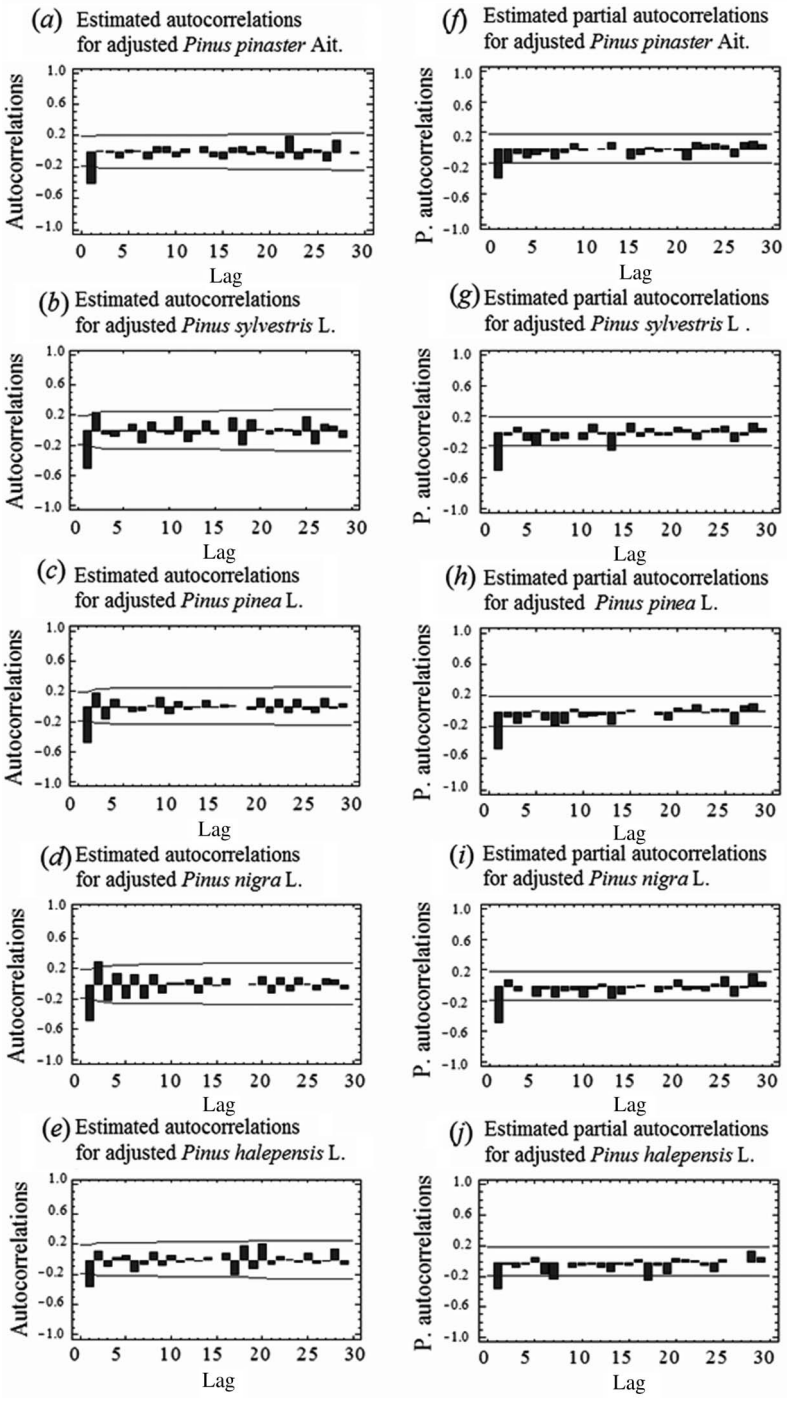


Figure 5. ACF and PACF of the MVC-NDVI time series. Dashed lines represent the 95% confidence interval.

Table 1. Magnitude of different errors made in the estimation of the preselected SARIMA(p,q,d) × (P,Q,D)$_S$ models (*Pinus pinaster* Ait.).

Preselected model	RMSE	MAE	MAPE	ME	MPE
(A) SARIMA(0,1,1) × $(0,1,1)_{37}$	0.0014	0.0385	12.8031	0.0020	–0.3140
(B) SARIMA(1,1,1) × $(0,1,1)_{37}$	0.0022	0.0409	13.5789	0.0028	–0.1240
(C) SARIMA(1,1,1) × $(1,1,1)_{37}$	0.0022	0.0419	13.8860	0.0027	–0.1750
(D) SARIMA(0,1,1) × $(1,1,1)_{37}$	0.0015	0.0393	12.9871	0.0019	–0.3422

RMSE, root mean squared error; MAE, mean absolute error; MAPE, mean absolute percentage error; ME, mean error; MPE, mean percentage error; SARIMA, seasonal autoregressive integrated of moving averages; p, autoregressive coefficient; q, moving average coefficient; d, differentiation time; P, seasonal autoregressive coefficient; Q, seasonal moving average coefficient; D, deterministic seasonal trend; S, seasonality.

Table 2. Results of five tests run on the residuals to determine whether each SARIMA(p,q,d) × (P,Q,D)$_S$ model is adequate for *Pinus pinaster* Ait. data.

Preselected model	RUNS	RUNM	AUTO	MEAN	VAR
(A) SARIMA(0,1,1) × $(0,1,1)_{37}$	OK	OK	*	OK	OK
(B) SARIMA(1,1,1) × $(0,1,1)_{37}$	OK	*	***	OK	OK
(C) SARIMA(1,1,1) × $(1,1,1)_{37}$	*	*	***	OK	OK
(D) SARIMA(0,1,1) × $(1,1,1)_{37}$	OK	OK	**	OK	OK

RUNS, test for excessive runs up and down; RUNM, test for excessive runs above and below median; AUTO, Box–Pierce test for excessive autocorrelation; MEAN, test for difference in mean first half to second half, VAR, test for difference in variance first half to second half; SARIMA, seasonal autoregressive integrated of moving averages; p, autoregressive coefficient; q, moving average coefficient; d, differentiation time; P, seasonal autoregressive coefficient; Q, seasonal moving average coefficient; D, deterministic seasonal trend; S, seasonality; OK, the model passes the test; *the model fails at the 95% confidence level; **the model fails at the 99% confidence level; ***the model fails at the 99.9% confidence level.

10 data values at the end of the time series were withheld to validate the model. Table 5 shows the errors of the model during the validation period. It can be seen in the same table that the models passed both the test for the suitability of the residuals and the test for randomness (p-value > 0.1).

Once the selected models were validated, they could be used for forecasting. Figure 6 presents simultaneously the observed values of the MVC-NDVI for each conifer species (points), the historic simulations (line) and the forecasts (line). The confidence interval limit was defined at 95%; this exigent confidence level means that the models are appropriate for making short-term forecasts. This type of forecasting is the most useful for monitoring systems in semi-real time.

Concerning the NDVI–climate relationships, the influence of both the temperature and precipitation level on the inter-annual evolution of MVC-NDVI time series was analysed. Observing the evolution of the NDVI with temperature, their relationship was shown to be direct. It even overlapped in some species; figure 7 displays this evolution for *Pinus sylvestris* L., considering the T_a. None of the calculated temperature variables contributed significantly to the SARIMA models. Significance was found, however, in the precipitation level variable for some species. Species subjected to summer water shortage, such as *Pinus nigra* L., *Pinus pinea* L. and *Pinus halepensis* L., showed a significant relationship with precipitation data, although only a very small contribution (table 6).

Table 3. Coefficients (C) and p-values associated with the selected SARIMA(p,q,d) × (P,Q,D)$_S$ models for each conifer species considered.

	Pinus sylvestris L. SARIMA(0,1,1) × $(1,1,1)_{37}$		*Pinus nigra* L. SARIMA(0,1,1) × $(1,1,1)_{37}$		*Pinus pinaster* Ait. SARIMA(0,1,1) × $(0,1,1)_{37}$		*Pinus pinea* L. SARIMA(0,1,1) × $(0,1,1)_{37}$		*Pinus halepensis* L. SARIMA(0,1,1) × $(0,1,1)_{37}$		*Juniperus thurifera* L. SARIMA(0,1,1) × $(0,1,1)_{37}$	
	C	p-value	C	p-value	C	p-value	C	p-value	C	p-value	C	p-value
MA(1)	0.6718	0.000	0.6288	0.000	0.6526	0.000	0.5691	0.000	0.4447	0.000	0.5824	0.000
SAR(1)	0.6486	0.000	0.6092	0.000								
SMA(1)	0.6448	0.000	0.7060	0.000	0.7683	0.000	0.7707	0.000	0.8020	0.000	0.7982	0.000

MA, non-seasonal moving average; SAR, seasonal autoregressive; SMA, seasonal moving average; SARIMA, seasonal autoregressive integrated of moving averages; p, autoregressive coefficient; q, moving average coefficient; d, differentiation time; P, seasonal autoregressive coefficient; Q, seasonal moving average coefficient; D, deterministic seasonal trend; S, seasonality.

Table 4. Errors made in the estimations of the selected SARIMA$(p,q,d) \times (P,Q,D)_S$ models.

	Pinus sylvestris L. SARIMA(0,1,1) $\times (1,1,1)_{37}$	*Pinus nigra* L. SARIMA(0,1,1) $\times (1,1,1)_{37}$	*Pinus pinaster* Ait. SARIMA(0,1,1) $\times (0,1,1)_{37}$	*Pinus pinea* L. SARIMA(0,1,1) $\times (0,1,1)_{37}$	*Pinus halepensis* L. SARIMA(0,1,1) $\times (0,1,1)_{37}$	*Juniperus thurifera* L. SARIMA(0,1,1) $\times (0,1,1)_{37}$
RMSE	0.0024	0.0014	0.0014	0.0012	0.0017	0.0013
MAE	0.0366	0.0309	0.0284	0.0275	0.0310	0.0278
MAPE	14.1398	13.2780	9.4029	10.7794		10.5006
ME	–0.0023	–0.0020	0.0028	–0.0010	–0.0003	0.0005
MPE	–5.3374	–3.1199	0.0667	–1.5640		–0.9395

RMSE, root mean squared error; MAE, mean absolute error; MAPE, mean absolute percentage error; ME, mean error; MPE, mean percentage error; SARIMA, seasonal autoregressive integrated of moving averages; p, autoregressive coefficient; q, moving average coefficient; d, differentiation time; P, seasonal autoregressive coefficient; Q, seasonal moving average coefficient; D, deterministic seasonal trend; S, seasonality.

Table 5. Forecast model statistics for the validation period.

	Pinus sylvestris L. SARIMA(0,1,1) × $(1,1,1)_{37}$	*Pinus nigra* L. SARIMA(0,1,1) × $(1,1,1)_{37}$	*Pinus pinaster* Ait. SARIMA(0,1,1) × $(0,1,1)_{37}$	*Pinus pinea* L. SARIMA(0,1,1) × $(0,1,1)_{37}$	*Pinus halepensis* L. SARIMA(0,1,1) × $(0,1,1)_{37}$	*Juniperus thurifera* L. SARIMA(0,1,1) × $(0,1,1)_{37}$
Statistics for validation period						
RMSE	0.0015	0.0007	0.0012	0.0007	0.0013	0.000
MAE	0.0328	0.0224	0.0311	0.0249	0.0296	0.0213
MAPE	9.1803	9.1551	10.8837	10.0353		9.8351
ME	0.0131	0.0022	0.0053	0.0061	–0.0087	–0.0031
MPE	2.6656	–0.7806	1.2274	1.6346		–3.0957
Tests for suitability of residuals						
RMSE		0.0377	0.0380	0.0350	0.0414	0.0368
RUNS	OK	*	OK	OK	OK	**
RUNM	OK	OK	OK	OK	OK	OK
AUTO	OK	OK	OK	OK	*	OK
MEAN	OK	OK	OK	OK	OK	OK
VAR	OK	OK	OK	OK	OK	OK
Tests for randomness of residuals						
RUNM	0.6779	0.4548	0.3609	0.5610	0.5610	0.5610
RUNS	0.2508	0.2508	0.9738	0.6697	0.4121	0.0115
Box–Pierce	0.3262	0.4916	0.4542	0.3476	0.1129	0.4049

RMSE, root mean squared error; MAE, mean absolute error; MAPE, mean absolute percentage error; ME, mean error; MPE, mean percentage error; RUNS, test for excessive runs up and down; RUNM, test for excessive runs above and below median; AUTO, Box–Pierce test for excessive autocorrelation; MEAN, test for difference in mean first half to second half, VAR, test for difference in variance first half to second half; OK, the model passes the test; *the model fails at the 95% confidence level; **the model fails at the 99% confidence level; ***the model fails at the 99.9% confidence level; SARIMA, seasonal autoregressive integrated of moving averages; p, autoregressive coefficient; q, moving average coefficient; d, differentiation time; P, seasonal autoregressive coefficient; Q, seasonal moving average coefficient; D, deterministic seasonal trend; S, seasonality.

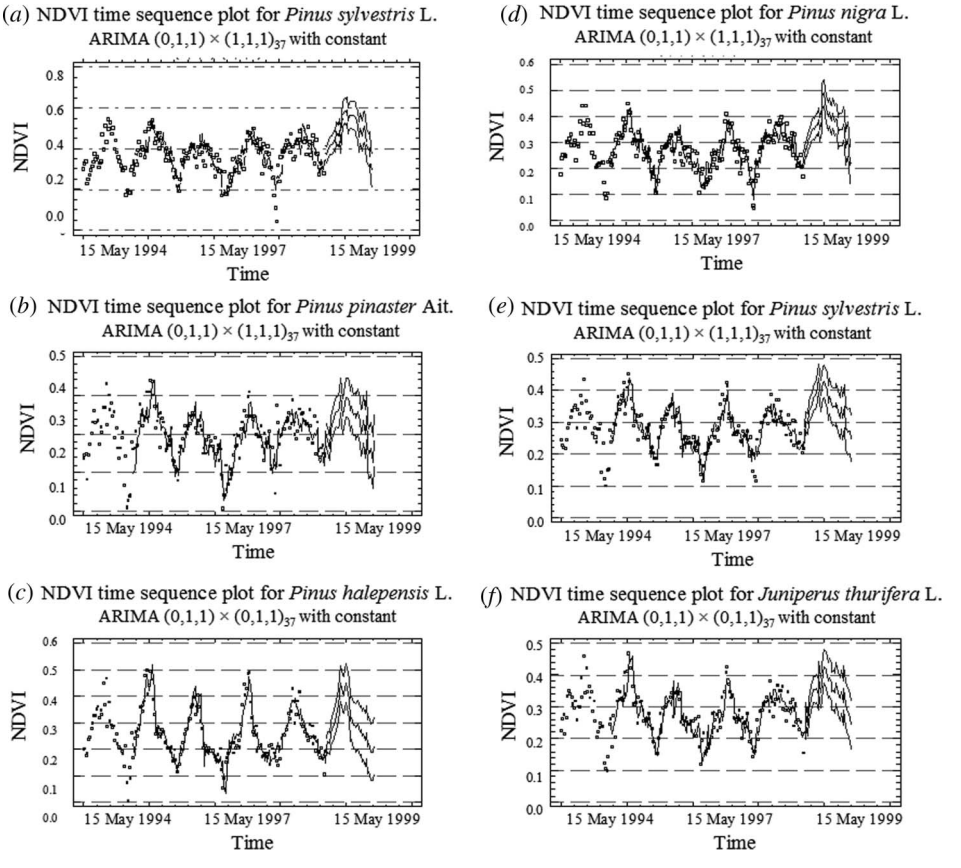


Figure 6. MVC-NDVI time sequence plots, showing the observed values of the MVC-NDVI for each conifer species (points), the historic simulations and the forecasts (lines).

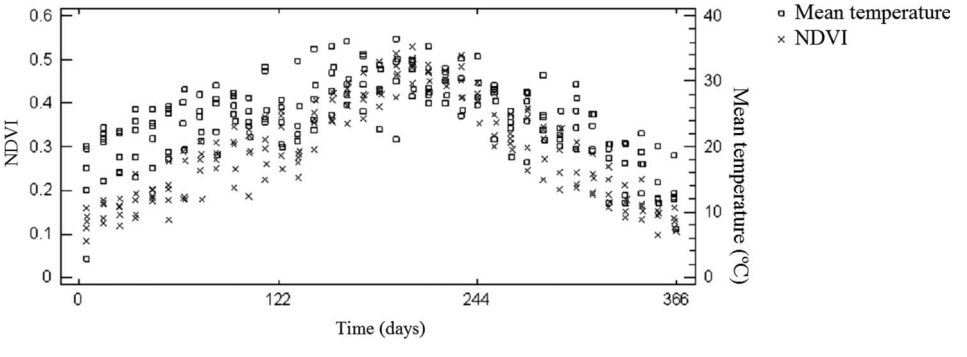


Figure 7. Inter-annual evolution of the MVC-NDVI and mean temperature (*Pinus sylvestris* L.).

Table 6. Coefficients (C) and p-values associated with selected SARIMA(p,q,d) × (P,Q,D)$_S$ models when climatic data were included.

	Pinus sylvestris L. SARIMA(0,1,1) × (1,1,1)$_{37}$		*Pinus nigra* L. SARIMA(0,1,1) × (1,1,1)$_{37}$		*Pinus pinaster* Ait. SARIMA(0,1,1) × (0,1,1)$_{37}$		*Pinus pinea* L. SARIMA(0,1,1) × (0,1,1)$_{37}$		*Pinus halepensis* L. SARIMA(0,1,1) × (0,1,1)$_{37}$		*Juniperus thurifera.* L. SARIMA(0,1,1) × (0,1,1)$_{37}$	
	C	p-value	C	p-value	C	p-value	C	p-value	C	p-value	C	p-value
MA(1)	0.6718	0.000	0.6079	0.0000	0.6526	0.000	0.5391	0.0000	0.4164	0.0000	0.5824	0.000
SAR(1)	0.6486	0.000	–0.5882	0.0000								
SMA(1)	0.6448	0.000	0.7122	0.0000	0.7683	0.000	0.7816	0.0000	0.8123	0.0000	0.7982	0.000
Precipitation level		–	–0.0002	0.0900		–	–0.0003	0.0100	–0.0002	0.0900		–
Mean temperature		–		–		–		–		–		–

MA, moving average; SAR, seasonal autoregressive ; SMA, seasonal moving average; –, without significance; SARIMA, seasonal autoregressive integrated of moving averages; p, autoregressive coefficient; q, moving average coefficient; d, differentiation time; P, seasonal autoregressive coefficient; Q, seasonal moving average coefficient; D, deterministic seasonal trend; S, seasonality.

As climate variables did not contribute significantly in the previously validated SARIMA models (SARIMA(0,1,1) × $(1,1,1)_{37}$ for *Pinus sylvestris* L. and *Pinus nigra* L. and SARIMA(0,1,1) × $(0,1,1)_{37}$ for the rest of the conifer species considered), they were used to calculate the predicted NDVI images for each conifer species and for each 10-day period. Using these models, and with the support of the Geographic Information System (GIS), images of the expected response were constructed.

A pilot study zone was selected in the southeast quadrant of Castile and Leon. The NDVI forecast was made for the different 10-day periods of 1997 on the basis of the models designed. Figure 8 compares the real and forecasted MVC-NDVI images classified in 0.05 NDVI unit intervals for the 10-day period between 30 June and 9 July 1997. Both images have the same MVC-NDVI values in five of the six conifer species considered; therefore it can be confirmed that the modelling performed allows shot-term forecasting. *Pinus pinaster* Ait., however, presented a lower value in the real NDVI image; this represents a natural risk situation that must be investigated in detail to understand the cause of this difference and to activate a warning.

Figure 9 shows the scatterplots of the MVC-NDVI predicted vs. the observed values for each conifer species considered in the pilot study area. There was good

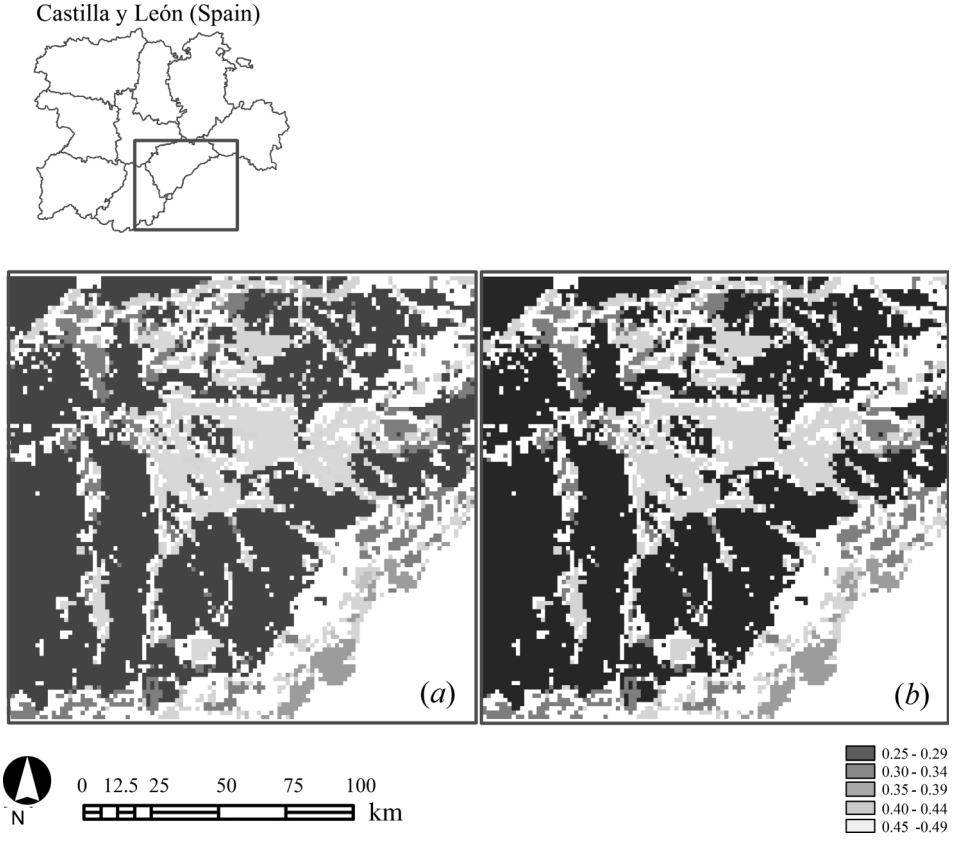


Figure 8. The 10-day period between 30 June and 9 July 1997. (*a*) real MVC-NDVI image; (*b*) forecasted NDVI image. The rectangle over the Castile and Leon perimeter represents the pilot study area.

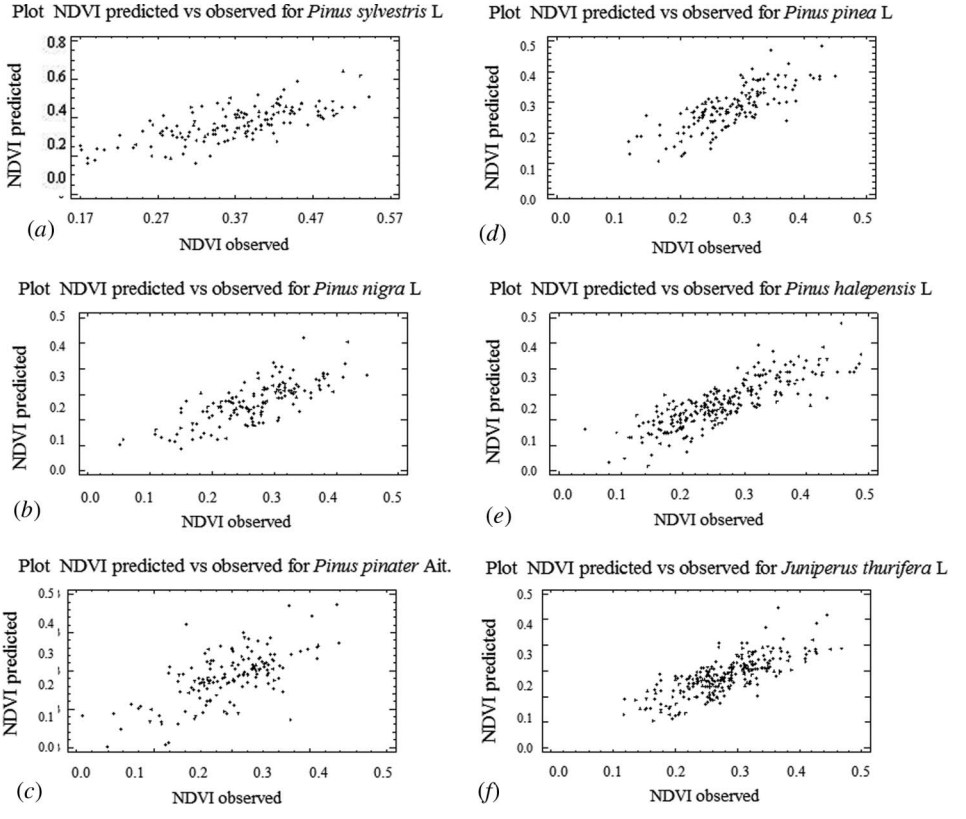


Figure 9. Scatterplots of the MVC-NDVI predicted vs. the observed values for each conifer species ((*a*)–(*f*)).

agreement between both values. R^2 varied from 0.67 (*Pinus pinea* L.) to 0.81 (*Pinus pinaster* Ait.). It should be noted that the spatial variability within each forest species was lost because the scatterplots related the MVC-NDVI predicted values to the averaged observed ones.

5. Discussion

Piwowar and Ledrew (2002) showed that time series modelling (based on ARMA analysis) of remote sensing imagery can be effectively used in analysing temporal sequences of Arctic sea-ice concentrations, stating that such new image analysis strategies could lead to measuring variability and to identifying changes in long sequences of remotely sensed imagery. Our study confirmed the use of ARMA modelling with remote sensing imagery. We used ARIMA models with MVC-NDVI time series to forecast the vegetation state and identify risk situations. An unexpected decrement in NDVI values is mainly related to natural risks such as wildfires, disease or insect defoliation. The model also led to the identification of the vegetation response to changes in growing conditions (mainly caused by precipitation and temperature variability). The selected SARIMA model allowed for an understanding of the cyclical structure of the temporal measurements, identifying a

moving-average regular term with a 10-day lag and an autoregressive 37 10-day period seasonal term with a one-season (1-year) component.

Concerning the NDVI–climate relationships, and according to Dye and Tucker (2003), Kaufmann *et al.* (2003) and Piao *et al.* (2003), this relationship depends, above all, on the climatology of the zone. Specifically, Deblonde and Cihlar (1993) stated that in cold zones the NDVI seems to be more correlated to the air temperature, so it is superimposed on the time series without contributing any significant information. The results of our study agree with this. None of the temperature variables considered contributed significantly to the SARIMA models previously validated for each species of conifer considered. Regarding the precipitation level, most studies that refer to the NDVI–precipitation relationship report that precipitation contributes important information to the interpretation of the temporal evolution of the NDVI (Yang and Merchand 1997, Wang *et al.* 2001, Holm *et al.* 2003, Li *et al.* 2004). In all these studies precipitation has a direct influence on the response of the NDVI and, as a consequence, on productivity. Our study, however, only agrees with these statements when species subjected to summer water stress were considered, and to a lesser degree. Nevertheless, vegetation response is related not only to climate but also to soil characteristics such as type, moisture and light availability. Additional research is needed to consider other possible sources of impact on the NDVI series.

Nevertheless, the use of data from sensors with higher temporal resolution could improve the forecasting process, allowing time series with shorter intervals. It is important to note, however, that the sensor needs to have a high stability over time to allow series of several years with the same characteristics. Nowadays, data series from the MODerate-resolution Imaging Spectroradiometer (MODIS) from both TERRA and AQUA platforms have been available since 2002 and could be used as inputs to the proposed methodology. In addition, data with a high spatial resolution would allow forecasting at lower scales. We leave these developments for future work.

Other vegetation indexes could also be used to describe the vegetation response from remote sensed data. McDonald *et al.* (1998) affirmed that vegetation indices should be selected according to the characteristics of the site under investigation. At low cover, the global environment monitoring index (GEMI) performed best, whereas at high cover, the soil adjusted vegetation index (SAVI) and the transformed soil adjusted vegetation index (TSAVI) were found to perform best. McDonald *et al.* (1998) concluded that further work is required to develop spectral indices specifically for conifer forests. However, Boegh *et al.* (2002), Huete *et al.* (2002) and Houborg and Soegaard (2004) reported the sensitivity of the enhanced vegetation index (EVI) to variations in leaf area index (LAI) for high biomass regions.

Lastly, the study of spatial variability also needs further research. MVC-NDVI values in the observed image are not spatially constant, although spatially constant values were considered in this study (specifically, the average MVC-NDVI values according to the species map).

6. Conclusions

ARIMA analysis of the time series of remotely sensed data is a potential tool for forecasting natural risks such as drought, fire risk or forest disease in forested areas. A vegetation monitoring system capable of forecasting the short-term response of forest vegetation using ARIMA models was designed to evaluate the risk associated with

unexpected decreases of the NDVI values. These models could be used as a basis for vegetation monitoring systems in semi-real time at a regional level.

Different SARIMA models that best suit the evolution of the NDVI series for each of the conifers species considered were defined. In the case of species subjected to climatic conditions, the most suitable models developed were mainly adjusted to a SARIMA$(0,1,1) \times (0,1,1)_{37}$ with a moving average component in both the non-seasonal and seasonal parts. In *Pinus sylvestris* L. and *Pinus nigra* L. an AR component appeared in the seasonal part: SARIMA$(0,1,1) \times (1,1,1)_{37}$. Using climatic variables (precipitation level) as regressors of the MVC-NDVI time series served to improve slightly the forecasting models when species subjected to summer water stress were considered. The models designed were used to elaborate a short-term forecast of the NDVI image for every one of the conifer species in each 10-day period.

References

Almeida, T.I.R., De Souza Filho, C.R. and Rossetto, R., 2006, ASTER and Landsat ETM+ images applied to sugarcane yield forecast. *International Journal of Remote Sensing*, **27**, pp. 4057–4069.

Aurenhammer, F., 1991, Voronoi diagrams – a survey of a fundamental geometric data structure. *ACM (Association for Computing Machinery) Computing Surveys*, **23**, pp. 345–405.

Bails, D.G. and Peppers, L.C., 1982, *Business Fluctuations: Forecasting Techniques and Applications* (Englewood Cliffs, NJ: Prentice-Hall).

Boegh, E., Soegaard, H., Broge, N., Hasager, C.B., Jensen, N.O. and Schelde, K., 2002, Airborne multispectral data for quantifying leaf area index, nitrogen concentration, and photosynthetic efficiency in agriculture. *Remote Sensing of Environment*, **81**, pp. 179–193.

Box, G.E.P. and Jenkins, G.M., 1976, *Time Series Analysis: Forecasting and Control* (San Francisco: Holden-Day).

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C., 1994, *Time Series Analysis Forecasting and Control* (New York: Prentice Hall).

Chang, T., 1973, *Reciprocal Distance Squared Method, A Computer Technique for Estimating Areal Precipitation* (Washington, D.C.: US Department of Agriculture, Agricultural Research Service).

Deblonde, G. and Cihlar, J., 1993, A multiyear analysis of the relationship between surface environmental variables and NDVI over the Canadian landmass. *Remote Sensing Reviews*, **7**, pp. 151–177.

Delgado, J.A., 1991, Clouds analysis and classification by using Meteosat images [in Spanish]. PhD thesis, University of Valladolid, Spain.

Dye, D.G. and Tucker, C.J., 2003, Seasonality and trends of snow-cover, vegetation index, and temperature in northern Eurasia. *Geophysical Research Letters*, **30**, pp. 1405–1412.

Farrar, T.J., Nicholson, S.E. and Lare, A.R., 1994, The influence of soil type on the relationships between NDVI, rainfall, and soil moisture in semiarid Botswana. II. NDVI response to soil moisture. *Remote Sensing of Environment*, **50**, pp. 121–133.

Gong, D. and Ho, C., 2003, Detection of large-scale climate signals in spring vegetation index (normalized difference vegetation index) over the Northern Hemisphere. *Journal of Geophysical Research*, **108**, pp. 4498–4510.

Holm, A.M.R., Cridland, S.W. and Roderinck, M.L., 2003, The use of time-integrated NOAA NDVI data and rainfall to assess landscape degradation in the arid shrubland of western Australia. *Remote Sensing of Environment*, **85**, pp. 145–158.

Houborg, R.M. and Soegaard, H., 2004, Regional simulation of ecosystem CO_2 and water vapor exchange for agricultural land using NOAA AVHRR and Terra MODIS satellite data. Application to Zealand, Denmark. *Remote Sensing of Environment*, **93**, pp. 150–167.

HUETE, A., DIDAN, K., MIURA, T., RODRIGUEZ, E.P., GAO, X. and FERREIRA, L.G., 2002, Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, **83**, pp. 195–213.

ICONA, 1990, *Second National Forest Inventory 1986–1995. Explanations and Methods* [in Spanish] (Madrid: ICONA (National Institute of Nature Conservation), Agriculture, Fisheries and Food Ministry).

ILLERA, P., FERNÁNDEZ, A. and CALLE, A., 1997, Operational forest fire danger indices derived from NOAA images. In *Integrated Applications for Risk Assessment and Disaster Prevention for the Mediterranean*, *16th EARSEL Symposium*, 20–23 May 1996, Malta, pp. 319–325 (Rotterdam: Balkema).

ILLERA, P., FERNÁNDEZ, A. and PÉREZ, A., 1995, A simple model for the calculation of global solar radiation using geostationary satellite data. *Atmospheric Research*, **39**, pp. 79–90.

JARLAN, L., TOURRE, Y.M., MOUGIN, E., PHILIPPON, N. and MAZZEGA, P., 2005, Dominant patterns of AVHRR NDVI interannual variability over the Sahel and linkages with key climate signals (1982–2003). *Geophysical Research Letters*, **32**, pp. 4701–4711.

JI, L. and PETTERS, A.J., 2004, Forecasting vegetation greenness with satellite and climate data. *IEEE Geoscience and Remote Sensing Letters*, **1**, pp. 3–6.

KAUFMAN, Y.J. and HOLBEN, P.N., 1993, Calibration of the AVHRR visible and near-IR bands by atmospheric scattering, ocean glint and desert reflection. *International Journal of Remote Sensing*, **15**, pp. 21–25.

KAUFMANN, R.K., ZHOU, L., MYNENI, R.B., TUCKER, C.J., SLAYBACK, D., SHABANOV, N.V. and PINZON, J., 2003, The effect of vegetation on surface temperature: a statistical analysis of NDVI and climate data. *Geophysical Research Letters*, **30**, pp. 2147–2159.

KOGAN, F.N., 1995, Application of vegetation index and brightness temperature for drought detection. *Advances in Space Research*, **15**, pp. 91–100.

KUMAR, M.R.R., 2004, Forecasting of onset of southwest monsoon over Kerala coast using satellite data. *IEEE Geoscience and Remote Sensing Letters*, **1**, pp. 265–267.

LI, J., LEWIS, J., ROWLAND, J., TAPPAN, G. and TIESZEN, L.L., 2004, Evaluation of land performance in Senegal using multi-temporal NDVI and rainfall series. *Journal of Arid Environments*, **59**, pp. 463–480.

LIU, W.T. and KOGAN, F.N., 1996, Monitoring regional drought using Vegetation Condition Index. *International Journal of Remote Sensing*, **17**, pp. 2761–2782.

LIU, W.T., MASSAMBANI, O. and NOBRE, C.A., 1994, Satellite recorded vegetation response to drought in Brazil. *International Journal of Climatology*, **14**, pp. 343–354.

LUTHER, J.E., FRANKLIN, S.E., HUDAK, J. and MEADES, J.P., 1997, Forecasting the susceptibility and vulnerability of balsam fir stands to insect defoliation with Landsat Thematic Mapper data. *Remote Sensing of Environment*, **59**, pp. 77–91.

MAKRIDAKIS, S., WHEELWRIGHT, S.C. and HYNDMAN, R.J., 1998, *Forecasting: Methods and Applications* (New York: John Wiley & Sons).

MCDONALD, A.J., GEMMELL, F.M. and LEWIS, P.E., 1998, Investigation of the utility of spectral vegetation indices for determining information on coniferous forests. *Remote Sensing of Environment*, **66**, pp. 250–272.

NAGLER, T., ROTT, H., MALCHER, P. and MÜLLER, F., 2008, Assimilation of meteorological and remote sensing data for snowmelt runoff forecasting. *Remote Sensing of Environment*, **112**, pp. 1408–1420.

PARUELO, J.M. and LAUENROTH, W.K., 1998, Interannual variability of NDVI and its relationship to climate for North American shrublands and grasslands. *Journal of Biogeography*, **25**, pp. 721–733.

PEÑA, D., TIAO, G.C. and TSAY, R.S., 2001, *A Course in Time Series Analysis* (New York: John Wiley & Sons).

PETERS, A., RUNDQUIST, D.C. and WILHITE, P.A., 1991, Satellite detection of the geographic core of the 1988 Nebraska drought. *Agricultural and Forest Meteorology*, **57**, pp. 35–47.

PIAO, S., FANG, J., ZHOU, L., GUO, Q., HENDERSON, M., JI, W., LI, Y. and TAO, S., 2003, Interannual variations of monthly and seasonal normalized difference vegetation index (NDVI) in China from 1982 to 1999. *Journal of Geophysical Research*, **108**, pp. 4401–4413.

PIWOWAR, J.M. and LEDREW, E.F., 2002, ARMA time series modeling of remote sensing imagery: a new approach for climate change studies. *International Journal of Remote Sensing*, **23**, pp. 5225–5248.

RIAÑO, D., MORENO, J.A., BARÓN, J. and USTIN, S.L., 2007, Burned area forecasting using past burned area records and Southern Oscillation Index for tropical Africa (1981–1999). *Remote Sensing of Environment*, **107**, pp. 571–581.

RICHARD, Y. and POCCARD, I., 1998, A statistical study of NDVI sensitivity to seasonal and interannual rainfall variations in southern Africa. *International Journal of Remote Sensing*, **19**, pp. 2907–2920.

SRIDHAR, V.N., DADHWAL, V.K., CHAUDHARI, K.N., SHARMA, R., BAIRAGI, G.D. and SHARMA, A.K., 1994, Wheat production forecasting for a predominantly unirrigated region in Madhya Pradesh (India). *International Journal of Remote Sensing*, **15**, pp. 1307–1316.

UNGANAI, L.S. and KOGAN, F.N., 1998, Drought monitoring and corn yield estimation in Southern Africa from AVHRR data. *Remote Sensing of Environment*, **63**, pp. 219–232.

WANG, J., PRICE, K.P. and RICH, P.M., 2001, Spatial patterns of NDVI in response to precipitation and temperature in the central great plains. *International Journal of Remote Sensing*, **22**, pp. 3827–3844.

YANG., Y.L. and MERCHAND, J.W., 1997, An assessment of AVHRR/NDVI–ecoclimatological relations in Nebraska, USA. *International Journal of Remote Sensing*, **18**, pp. 2161–2180.

ZHOU, L., TUCKER, C.J., KAUFMANN, R.K., SLAYBACK, D., SHABANOV, N.V. and MYNENI, R.B., 2001, Variations in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999. *Journal of Geophysical Research*, **106**, pp. 20069–20084.