MODELLING NDVI TIME SERIES TO FILL GAPS OF METEOROLOGICAL DATA

Margarita González Loyarte[1], Massimo Menenti[2] , Ángela Magdalena Diblasi[3,4]

1. Instituto Argentino de Investigaciones de las Zonas Áridas (IADIZA/ CONICET CCT Mendoza), Mendoza, Argentina; gloyarte@mendoza-conicet.gob.ar
2. University of Delft, Department of Geoscience and Remote sensing, Delft, The Netherlands; M.Menenti@tudelft.nl
3. Facultad de Ciencias Económicas, Univ. Nacional de Cuyo, Mendoza, Argentina; angelad@raiz.uncu.edu.ar
4. Área de Ciencias Exactas,CONICET-CCT, Mendoza, Argentina; adiblasi@mendoza-conicet.gob.ar

ABSTRACT

We applied a measure of foliar phenology to interpolate climate statistics and produce a bioclimate classification for a vast plain in Argentina, with sparse weather observations. As a measure of foliar phenology, we used parameters obtained by modelling NDVI time series with a Fast Fourier Transform (FFT) applied to a 9-year time series of NOAA-AVHRR NDVI GAC images. FFT decomposes the series into an average signal and two sinusoidal components. Selected FFT parameters were mean NDVI, amplitude and phase for a 1-year period. Climate data were annual rainfall (P) and mean temperature (T) expressed as Potential Evapotranspiration (PET) estimated by an empirical equation (PET= 68.64 T). P/PET ratio was related to FFT parameters by fitting a multiple linear regression model with P/PET as predicted variable and FFT parameters as predictive variables. The regression model, that explained 92% of the P/PET variation, was then applied to the entire selected images (parameters) to obtain a map of the P/PET ratio. The P/PET map was compared with existing climate maps and ancillary data to derive a consistent bioclimate map. Mean annual phenological rhythm was graphed for each bioclimate by reconstructing the yearly NDVI curve. This shows that aridity reduces the contrast between minimum and maximum NDVI and that time of maximal vegetation cover varies from January (semiarid) to April (subdesert). The proposed method is an adequate tool to extend meteorological data into regions where climate data have an uneven coverage or poor spatial resolution. This finding shows that each selected FFT parameter was necessary because none was significant by itself.

INTRODUCTION

In extensive arid and semiarid regions, very often there is a gap of meteorological data. NDVI time series allow indirect detection of climate conditions: Normalized Difference Vegetation Index (NDVI = (NIR-R)/(NIR+R) images are strongly related to photosynthetic active radiation (PAR) (1) and consequently to the activity of vegetation cover, expressed as foliar phenology (2) and its regional changes with climate conditions (3). The response of NDVI data to rainfall has been largely studied (4). The strong linear relationship when annual rainfall ranges from 150 to 1000 mm (5) allows detection of drought and 'wetter' conditions on Sahelian vegetation for the period 1981-2003 (6). The usefulness of NOAA-AVHRR NDVI series in the study of interannual variability produced by ENSO events has been demonstrated (7,8). Moreover, in places where rainfall data are sparse, the use of NDVI series instead of rainfall data improved the correlations with ENSO indices predicting drought onset in Northeastern Brazil four months in advance with 68% success (7).

This ability of time series of NDVI to express climate conditions has proved to be synthesized by Fourier parameters (9,10,11). The Fast Fourier Transform algorithm allows decomposing, for every

pixel, the temporal profile of NDVI series in an average signal plus N/2 sinusoidal components, with N being the length of the time series expressed as number of images (9, 10). The average signal is the mean NDVI for the whole time series, and periodic (sinusoidal) components are characterized by amplitude and phase and are associated with a given period, e.g. 9, 4.5, 3, 1.5, 1, 0.5 years. The amplitude represents a measure of the maximum variability of NDVI at a given period, and phase is the time lag between this maximum and the initial point of the series. The simplicity of NDVI Fourier parameters allows relating and understanding the effect of rainfall and temperature on vegetation units (12) and foliar seasonality (13) in Southern Africa; the mean NDVI and amplitude for a one-year period as sensitive indicators of climate variability (11); the usefulness of phase image to distinguish spatial and seasonal rainfall variations (14) and of the mean NDVI to detect rainfall regimes and climate types (7) in Northeastern Brazil.

The objective of this paper is to synthesize the methodology applied to a NOAA-AVHRR NDVI GAC series of monthly data for 9 growth cycles to fill gaps of meteorological data generating a bioclimate map through the P/PET ratio, for a vast plain of Central Western Argentina (15,16).

METHODS

Study area

The study area is the eastern plain of Mendoza (Argentina) that extends N-S for 440 km and is 100-150 km wide (W-E). It is a sandy plain covered with dryland vegetation: open *Prosopis* woodland and shrub steppe of *Larrea* spp., *Atriplex lampa*, etc. (17); 12% of the plain is irrigated for fruit trees, vineyards and orchards. Only 14 meteorological stations cover the area and over different periods (10 to 55 years). Mean annual temperature varies N-S from 18.7 ºC to 15 ºC. Annual rainfall, with high interannual variability (35-67%), increases eastward and southward from an average of 155 mm in the north to an average of 403 mm in the south; 78% of rains are concentrated in spring-summer.

Data, methods and procedures

Satellite data and image processing

Satellite data is a series of 9-year monthly NOAA-AVHRR NDVI GAC images (7.6 x 7.6 km), from July 1982 to June 1991. The series of NDVI images was processed in the Netherlands. There, at the National Aerospace Laboratory (NLR) a special so-called Mixed Radix Fast Fourier Transform algorithm (FFT) has been developed for the processing of time series of images (9,18,19). N is the length of the time series (108 images) and can be factored in the radix numbers 2, 3, 4 and 5 and, as the samples are equidistant in time (one month), the FFT algorithm allows decomposing, for every pixel, the NDVI series into an average signal plus the N/2 sinusoidal components. The resulting images were: mean NDVI for the whole series and the periodic components, i.e. amplitude and phase for 9, 4.5, 3, 1.5, 1 years, and 6, 4, 3, 2 months. Among them, the most sensitive bands were selected.

Selection of the most relevant FFT parameters (bands) was made according to the amplitude variance contribution of each period to total amplitude variance. Total variance of the time series was calculated as the sum of the variances of the individual terms as follows (20):

$$\text{Total variance} = \sum_{l=1}^{n} \frac{amplitude_j^2}{2} \quad (1)$$

where *j* is each term (period) in the series and *n* is the total number of terms. The relative contribution of each term is computed by dividing the individual variance for each term by the total variance. Contribution was calculated only for the periods: 9, 4.5, 3, 1.5,1 years and 6 months.

Meteorological data

Meteorological data used were mean annual rainfall (P) and mean annual temperature (T). The P/PET bioclimatic index was applied because it belongs to the synthetic indices that consider the climate requirements of vegetation. PET, potential evapotranspiration, i.e. the evaporative demand of the atmosphere over plants during a growing cycle, was estimated by the empirical equation PET = 68.64T; with T being mean annual temperature in ^{0}C (21). For oases, the average irrigation of 800 mm was added to mean annual rainfall to consider actual water availability.

The limits of bioclimate zones were adapted from Le Houérou (21,22,23) as follows: P/PET 0.06-0.15 as bioclimate class subdesert; 0.16-0.24 lower arid; 0.25-0.33 upper arid; 0.34-0.41 lower semiarid; 0.42-0.50 upper semiarid; 0.51-0.59 lower subhumid; 0.60-0.68 mid subhumid ; 0.69-0.75 upper subhumid; ≥ 0.76 humid.

Regression model

For each meteorological station, the closest pixel was selected and the digital numbers (DN) extracted for the selected bands to carry out a multiple linear regression analysis. The stations in the oases served to calculate P/PET for both oases and neighbouring drylands. On the borders of the oases, the P/PET index was obtained by averaging the indices with and without irrigation. In this way, and based on 14 meteorological stations, 27 values of the P/PET index were obtained. A multiple linear regression model was fitted between the selected FFT parameters (independent variables) and the P/PET (dependent variable) from 27 data. This model was expected to predict the P/PET bioclimatic index. "R" statistical free software was used (http://www.r-project.org).

On a first step a regression model with an *Indbi* response variable (P/PET index) and explanatory variables: *mean NDVI*, *amp*1 (amplitude for a one-year period), and *phase*1 (phase for a one-year period) was considered (24). Given the lack of normality in the response variable, it was transformed using the Box-Cox procedure (25) to: $Indbit=(Indbi^{\lambda}-1)/\lambda$, where the value for parameter λ was estimated as *0.4734707*. Then, a regression model was fitted under the assumptions of normality, independence, and homoscedasticity for the errors ε_i, i=1,….,27, verified through the analysis of residuals.

Generation of a map of the P/PET bioclimatic index

The coefficients of the obtained linear regression model were applied pixel by pixel to all three bands, mean NDVI, amplitude and phase for a one-year period as X_1, X_2 and X_3 respectively, generating the P/PET image where DNs are the bioclimatic index values. The P/PET index image was re-codified according to previous limits.The analysis of ancillary data (climate maps, climate data, geomorphology, land use, vegetation, etc.) drove to an adjustment of the P/PET classes to obtain the proposed bioclimate map. Phenological pattern was modelled for each bioclimate class.

Phenological pattern (monthly NDVI curve)

Once the time series is decomposed into its periodic components, the inverse process can be performed modelling the monthly NDVI for a year with only those periodic components conveying the most information disregarding noise. The resulting modelled curve shows the rhythm of phenology and we call it phenological pattern (26).The modelled annual NDVI curve, I(*t*), was calculated using the following equation (12):

$$I(t) = \sum_{n=1}^{N} \left[A_n * \cos\left(\omega_n t + \varphi_n\right) \right]$$ where: *I(t)* = reconstructed time series (with t= 1,…12 as first year of the series); A_n = amplitude value; φ_n = phase value *in radians; n* = index indicating periods; $\omega = 2\pi/n$ = frequency; we have considered the frequencies corresponding to periods: 9, 4.5, 3 and 1 years, in addition to frequency 0 whose amplitude is the mean NDVI value; *t* = time, 1 to 12 months (from July till June). Thus, monthly NDVI was calculated and graphed without the noise introduced by very brief periods (4,3,2 months) and without periods like 2.25, 1.8, 1.5 years, etc., with no clear phenologic information and little contribution to the observed NDVI signal. The contribution of the six-month period was negligible for the plain and was not used (15). Each curve corresponds to one pixel of the represented bioclimate class.

RESULTS

Fast Fourier Transform components

One hundred and eight images were obtained from applying the FFT algorithm. The one-year period contributed the most to total amplitude variance (43 to 86%), except for the northern arid plain (19-22%), expressing a single growing cycle a year. Here, the contribution of the 9-year period was dominant (52 to 56%) owing to high rainfall variability reflecting rare events only once over the 9 cycles. Contribution for the 4.5 (1-17%) and 3-year (2-15%) was higher than for the 6-month (0-3%). Therefore, only the one-year period was kept.

Selected FFT parameters for regression analysis were mean NDVI, amplitude and phase for the one-year period (Figure 1). The mean NDVI and amplitude images look similar, with the highest value for the irrigated oases; natural vegetation shows a N-S humidity gradient. The phase image contrasts with the other two, showing low values (short phases) for the oases, which expresses that maximum NDVI occurs early in the season, and high values (long phases) for the northern plain, where maximum NDVI occurs late in the season.

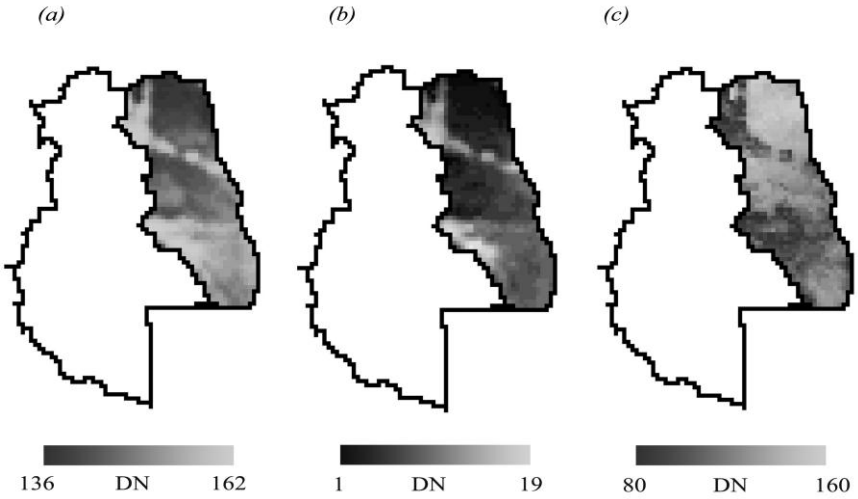

Figure 1: Images of Fourier coefficients expressed in digital numbers (DN): (a) mean NDVI, (b) amplitude and (c) phase for the one-year period obtained from the monthly NOAA-AVHRR NDVI GAC series July 1982 – June 1991 for the plain of Mendoza Province (Argentina).

Regression model

The multilinear regression model to predict the bioclimatic index (*Indbit*) showed a proper fit to the observed data (*p-value:* 2.510 x 10^{-13} for the goodness-of-fit *F*-test with a statistic of *99.7* on *3* and *23 DF*). This led to an adjusted Multiple *R*-Squared of *0.9193*. The residual standard error was 1.018 on 23 *DF*. No single variable was significant by itself (all p-values > 0.10) in the fitted model, *92%* of the total response variation (P/PET) is explained when using all three Fourier parameters together. Then, the regression model fitted to data result:

$$\hat{E}(Indbit_i) = -25.77468 + 0.27077\,meanNDVI_i + 0.29754\,amp1_i - 0.05410\,phase1_i \quad I=1,\ldots,27$$

Generation of a P/PET ratio map (bioclimatic index)

By applying the *Indbit* model to the whole plain, continuous bioclimate information was obtained (Figure 2a). Recoding this image according to adapted limits of Le Houérou's bioclimatic index led to a raw bioclimatic map where edaphic influence (irrigated oases, dunes, etc.) masked the influence of climate (Figure 2b). This result was clarified with the support of ancillary data and through a new recodification and editing, thus obtaining the proposed bioclimate map (Figure 2c).

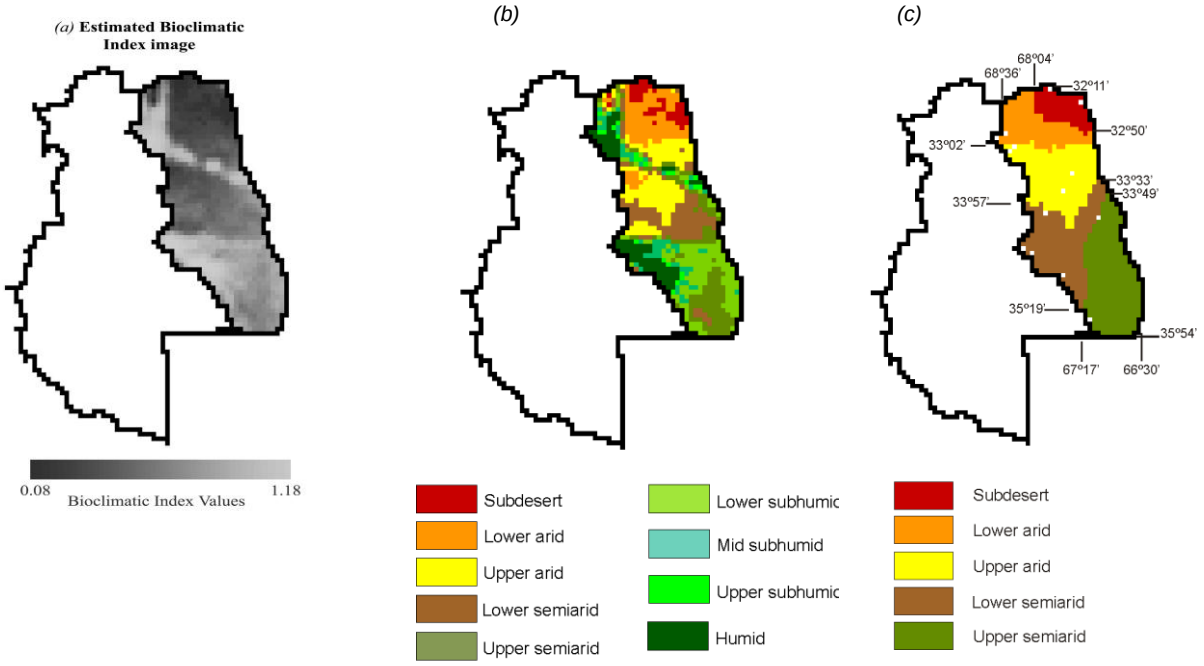

Figure 2: Bioclimate map derived from linear regression. (a) Bioclimatic image estimated from a multiple linear regression model (adjusted multiple R^2 of 0.9193, p-value $<10^{-12}$). (b) Derived bioclimate map. (c) Final adjustment according to ancillary data, bioclimate map proposed. White dots show the location of meteorological stations (Adapted from 15, 16).

Phenological pattern (monthly NDVI curve)

For each of the five bioclimate classes the modelled NDVI characterizes the influence of climate on the phenological variation throughout a growing cycle.

CONCLUSIONS

Linear regression with FFT parameters

Through the multilinear regression model to predict the P/PET index the three Fourier parameters together explains 92% of the P/PET variation; none of them conveys enough information by itself. Resutls of this model are comparable to the regression coefficients (0.80 to 0.93) obtained by (12) for amplitude at one-year and 6-month periods with the Budyko Aridity index in Southern Africa and with (27) who explained 70-80% of the spatial variability in NDVI seasonal extremes for different plant functional types through climate indices based on temperature and rainfall.

Fourier parameters are independent, give complementary information and describe the NDVI series in a simple and clear manner. The independence of Fourier parameters helps describe climate aspects like spatial and seasonal rainfall variation expressed by phase in Northeastern Brazil (14), whereas mean NDVI may express rainfall regimes and climate types (7). Also, mean NDVI and amplitude for a one-year period were found to be sensitive indicators of climate variability (11). Amplitude at a one-year period gives information on the contrast of vegetation cover along the growing cycle; the maximum contrast is in the irrigated oases where water is not scarce and the majority of permanent crops are deciduous. The lowest one-year amplitude is related to aridity and low plant cover, like in the north of the plain with many species with persistent, coriaceous and small leaves. Here, amplitude at the 9-year period presents higher values than at the one year period and expresses the interannual variability of rainfall. These results are coincident with those for Southern Africa where it was found that amplitude at 9 and 4.5 years increased as aridity increased (10). High 9-year amplitude was also found for the Acacia woodland-bushland of the Kalahari Desert, consistently with high interannual coefficients of variation in both rainfall and NDVI calculated for individual sites (13).

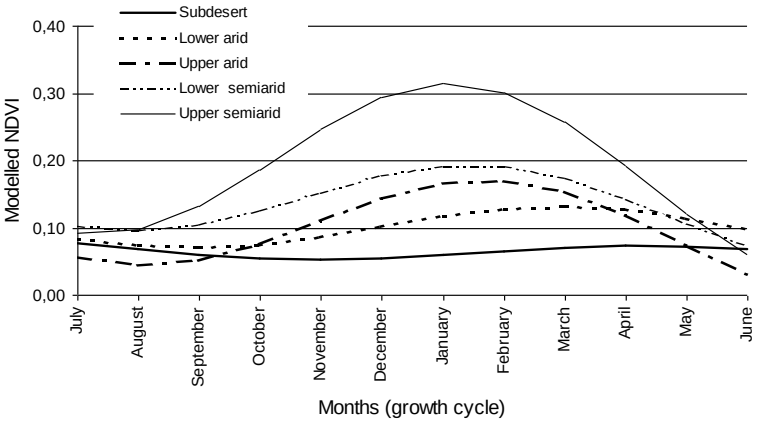

Figure 3: Modelled yearly NDVI curve (phenological pattern) for the bioclimate classes proposed for the eastern plain of Mendoza, Argentina (Adapted from 16).

Bioclimate map

Through this regression model, continuous bioclimate information on the plain was produced this one needed adjustment according to ancillary data to correct for edaphic influences. Here it is clear that sudden vegetation cover changes may yield confusing information on bioclimate, such as irrigated oasis where water supply transforms arid and semiarid environments, assigning them a humid or subhumid bioclimate. The influence of dunes lowers the climate class like in the 'island' in the SE of the southern plain, which responds to grasslands on gentle dunes. Detailed analysis and criteria applied in each case is found in (16). The consistency of the proposed bioclimate map is documented by the spatial continuity of classes and their geographical distribution and by its coherence with existing climate maps.

Phenological pattern

Modelling the NDVI pattern keeping the meaningful frequencies that express the intraannual variation at one-year period, and eventually 6-month period like in South Africa, and interannual variations (periods like 9, 4.5, 3 years) without the noise of short periods (high frequencies) help to understand, describe and compare the average vegetation pattern along areas with different climates (12).

The model of annual NDVI rhythm shows a progression in the maximum NDVI (amplitude) and the time it occurs (phase), between subdesert and upper semiarid climate (Figure 3). The amplitude increases with increase in humidity and phase is shorted by higher rainfall, varying from January (semiarid climate) to April (subdesert climate). Thus, each climate condition has a different vegetative expression throughout the year. In our case, there is one cycle per year, since the 6-month component showed negligible influence on intra-annual variability in NDVI.

This proposal allows expanding the climate knowledge from weather stations to the entire plain through the phenological expression of vegetation (in natural conditions) for this reflects climate conditions. The proposed bioclimate classes show a continuous spatial distribution, evidence of spatial coherence and class homogeneity, characteristics that reflect the influence of climate. When dealing with NDVI as a climate expression, management of vegetation cover (afforestation, fires, cultivation, etc.) and geomorphic features (dunes) should always be considered because they may modify the climate pattern derived from NDVI series.

ACKNOWLEDGEMENTS

The authors are grateful to Nelly Horak for her assistance with the english text, and to María Cecilia Scoones and Viviana Lotfi W.for their assistance with figures.

REFERENCES

1 Choudhury B J, 1987. Relationships between vegetation indices, radiation absorption, and net photosynthesis evaluated by a sensitive analysis. Remote Sensing of Environment, 22: 209-233.

2 Tucker C J & P J Sellers, 1986. Satellite remote sensing of primary production. International Journal of Remote Sensing, 7: 1395–1416.

3 Townshend, J R G & C O Justice, 1986, Analysis of the dynamics of African vegetation using the normalized difference vegetation index. International Journal of Remote Sensing, 7: 1435-1445.

4 Hielkema J U, S D Prince & W L Astle, 1986, Rainfall and vegetation monitoring in the Savanna Zone of the Democratic Republic of Sudan using the NOAA Advanced Very High Resolution Radiometer. International Journal of Remote Sensing, 7: 1499-1513.

5 Malo A R & S E Nicholson, 1990, A study of rainfall and vegetation dynamics in the African Sahel using normalized difference vegetation index. Jornal of Arid Environments, 19: 1-24.

6 Anyamba A & C J Tucker, 2005, Analysis of Sahelian vegetation dynamics using NOAA-AVHRR NDVI data from 1981-2003. Journal of Arid Environments, 63: 596-614.

7 Liu WT & R I Negrón Juárez, 2001. ENSO drought onset prediction in northeast Brazil using NDVI. International Journal of Remote Sensing, 22: 3483-3501.

8 Gurgel HC & N J Ferreira, 2003. Annual and interannual variability of NDVI in Brazil and its connections with climate. International Journal of Remote Sensing, 24: 3595-3609.

9 Menenti M, S Azzali, W Verhoef & R Van Swol, 1993. Mapping agroecological zones and time lag in vegetation growth by means of Fourier analysis of time series of NDVI images. Advances in Space Research, 13: 233-237.

10 Azzali S & M Menenti (eds.), 1996, Fourier analysis of temporal NDVI in the Southern African and American continents (Wageningen, The Netherlands: Winand Staring Centre for Integrated Land, Soil and Water Research, Report 108) 168 pp.

11 Roerink G J, M Menenti, W Soepboer & Z W Su, 2003. Assessment of climate impact on vegetation dynamics by using remote sensing. Physics and Chemistry of the Earth, Parts A/B/C**,** 28: 103-109.

12 Azzali S, & M Menenti, 2000. Mapping vegetation-soil-climate complexes in southern Africa using temporal Fourier analysis of NOAA-AVHRR data. International Journal of Remote Sensing, 21: 973-996.

13 Fuller D O & S D Prince, 1996. Regional-scale foliar phenology in tropical Southern Africa: An application of the Fast Fourier Transform to time series of satellite imagery. In Fourier analysis of temporal NDVI in the Southern African and American continents, edited by S. Azzali & Menenti (Winand Staring Centre for Integrated Land, Soil and Water Research, Report 108: Wageningen, The Netherlands) 113-132.

14 Negrón Juárez R I & W T Liu, 2001. FFT analysis on NDVI annual cycle and climatic regionality in Northeast Brazil. International Journal of Climatology**,** 21: 1803–1820.

15 González Loyarte M M, M Menenti & A M Diblasi, 2008. Modelling bioclimate by means of Fourier analysis of NOAA-AVHRR/NDVI time series in western Argentina. International Journal of Climatology, 28: 1175-1188.

16 González Loyarte M M, M Menenti & A M Diblasi, 2009. Mapa bioclimático para las Travesías de Mendoza (Argentina) basado en la fenología foliar. Revista de la Facultad de Ciencias Agrarias, XLI: 105-122.

17 Roig F A, M M González Loyarte, E Martínez Carretero, A Berra & C Wuilloud, 1992. La Travesía de Guanacache, Tierra Forestal. Multequina, 1: 83-91.

18 Menenti M, S Azzali, W Verhoef & R Van Swol, 1991. Mapping agroecological zones and time lag in vegetation growth by means of Fourier analysis of time series of NDVI images. Monitoring Agroecological Resources with Remote Sensing and Simulation (MARS). (Winand Staring Centre for Integrated Land, Soil and Water Research, Report 32: Wageningen, The Netherlands) 46 pp.

19 Verhoef W, 1996. Application of Harmonic Analysis of NDVI Time Series (HANTS). In Fourier analysis of temporal NDVI in the Southern African and American continents, edited by S. Azzali & Menenti (Winand Staring Centre for Integrated Land, Soil and Water Research, Report 108: Wageningen, The Netherlands) 19-24.

20 Jakubauskas M E, D R Legates & J H Kastens, 2001. Harmonic analysis of time-series AVHRR NDVI data. Photogrammetric Engineering & Remote Sensing, 67: 461-470.

21 Le Houérou H N, 1989.Classification écoclimatique des zones arides (s.l.) de l'Afrique du Nord. Ecologia Mediterranea, 15: 95-144.

22 Le Houérou H N, 1999. Estudios e investigaciones ecológicas de las zonas áridas y semiáridas de Argentina (Instituto Argentino de Investigaciones de las Zonas Áridas: Mendoza, Argentina) 228 pp.

23 Le Houérou H N, 2004. An Agro-Bioclimatic Classification of Arid and Semiarid Lands in the Isoclimatic Mediterranean Zones. Arid Land Research and Management, 18: 301 –346.

24 Faraway J, 2004. Linear Models in R. In: Texts in Statistical Science (Chapman & Hall/ CRS, London) 11-71.

25 Box G E P & D R Cox,. 1964. An analysis of transformations. Journal of Royal Statistics Society Series B, 26: 211-246.

26 González Loyarte M M, M. Menenti & F A Roig, 2010. Patrones fenológicos de la Provincia de Mendoza, Argentina, mediante serie temporal de imágenes NOAA-AVHRR NDVI GAC. Boletín de la Sociedad Argentina de Botánica, 45: 343-362.

27 Potter C S & V Brooks, 1998. Global analysis of empirical relations between annual climate and seasonality of NDVI. International Journal of Remote Sensing,19: 2921-2948.