Analysis of bioclimatic time series and their neural network-based classification to characterize drought risk patterns in South Italy

G. Incerti[a], E. Feoli[b], L. Salvati[c] A. Brunetti[c]

[a] Dipartimento di Biologia – Università di Trieste – via L. Giorgieri, 10 – I34127 Trieste, Italy

[b] Dipartimento di Biologia – Università di Trieste – via E. Weiss, 2 – I34127 Trieste, Italy

[c] UCEA – Ufficio Centrale di Ecologia Agraria – via del Caravita, 7/a – I00186 Roma, Italy

Corresponding Author:

dr. Guido Incerti

incerti@univ.trieste.it

Tel. +39 040 558 3889

Fax: +39 040 568855

Abstract

A new approach to characterize geographical areas with a drought risk index (DRI) is suggested with an application of an artificial neural network classifier (ANN). It computes a DRI based on a value equal to 1 assigned to situations of maximal drought risk and a value equal to -1 assigned to the situation with the lowest drought risk. The two situations are defined by integrals of a set of bioclimatic variables (rainfall, min temperature, max temperature, soil available water content, potential and real evapo-transpiration and Normalized Difference Vegetation Index) in a set of temporal units called Operational temporal Units (OtU). The OtUs are moving time windows with width (length) equal to the period defined for each variable by the Fast Fourier Transform method (FFT). 600 sites were sampled from 4 regions of Italy and the data were stored into a GIS database. It is suggested that for each of the area the graph showing the behaviour of DRI in these last 12 years, could be considered as a "finger print for drought" of the area. The classification of the areas based on the DRI graphs showed that, at the scale used in this work, DRI has no dependence from land cover class, but rather is related to the geographic macro area where the OGUs are located. The method can be used for indirect prediction of drought in given geographical areas if it is assumed that the areas showing high DRI for long time are the most prone to drought.

Key words: Drought, NDVI, Time series analysis, Fourier spectral analysis, Artificial Neural Network.

Introduction

According to its commonest interpretation, drought is the consequence of a negative persistent rainfall anomaly, often concomitant with other meteorological or pedologic events (soil erosion, high air temperature, high wind speed), regarding the average pattern to which the biotic component of the ecosystem is adapted in a given region (Wilhite, 1992).

In past decades, since the AI (Aridity Index) by Thornthwaite (1948), given by the ratio between annual precipitation and potential evapo-transpiration and devoted to the classification of moist climates, several drought indices have been proposed. Keetch and Byram (1968) designed a drought index specifically for fire potential assessment, which is a number representing the net effect of evapo-transpiration and rainfall in producing cumulative moisture deficiency in soil layers. The Palmer Drought Severity Index (Palmer, 1965) uses temperature and rainfall, as well as soil local Available Water Content (AWC) in a formula to determine dryness, and has been adopted as the official tool in drought relief programs by many U.S. government agencies (Hayes,

2003). McKee *et al.* (1993) developed the Standardized Precipitation Index (SPI), which only requires data for rainfall, based on the concept that rainfall deficits over varying periods or timescales influence ground water, reservoir storage, soil moisture, snowpack, and streamflow.

An index integrating climatic and remotely sensed information has been proposed by Feoli *et al*. (2003). It is based on the assumption that a desertification risk should be high when the yearly average temperature is high all over the year, the yearly amount of rainfall is concentrated only in some periods of the year and the NDVI after the dry season is low. Recently there is a tendency to evaluate drought severity using several indices or variables (Wu and Wilhite, 2004). Wilhite (2000) pointed out that it is important to use appropriate and reliable drought indices in decision-making, and that consulting more than one index is necessary and critical. To better track and assess the severity of droughts, in U.S. an integrated Drought Monitor product has been developed (Svoboda *et al.* 2002).

Following Tadesse *et al*. (2004), a large historical data set is necessary to identify relationships between the different parameters involved in determining drought, and to distinguish time patterns that could be used to predict drought. In this respect, it is critical to have an efficient tool to extract information from large databases and to deliver relevant information for drought risk management.

The present paper is in line with these suggestions and it proposes a new approach to characterize geographical areas by a drought risk index (DRI) based on the application of an artificial neural network classifier (ANN) that considers several bioclimatic variables at the same time without combining them mathematically in an index as it is usually done, but that is able to measure how much a situation is approaching the relative maximum risk of drought .

Artificial neural networks provide a methodology that has been used in different areas of research, such as computational economics, medical systems, molecular biology and ecology; the neural network application has recently increased in many different modelling studies, where time series analysis and prediction are a major goal (e.g. Moshiri, 1999; Kwak and Changwon, 1997; Henderson, 2000; Gopal *et al*., 1999). ANNs are appropriate tools in conditions where the relationships between the variables do not follow a formally defined model; nevertheless with ANNs it is possible to predict the future pattern of the system by extracting knowledge from the past. ANNs have been proved to be a powerful method in prediction, exceeding traditional linear modelling methods (Laepes and Farben, 1987). Its capabilities are based on the following features: a) ANNs make no assumptions about the nature of the data distribution and are not therefore, biased in their analysis (White, 1992); b) ANNs are considered as the best method to estimate non-linear relationships (Wasserman,

1989; Hoptroff, 1993); c) ANNs perform well with missing or incomplete data (Kuo and Reitsch, 1995).

Since the first half of 1990's, the US National Aeronautics and Space Administration (NASA), National Oceanic and Atmospheric Administration (NOAA), and US Geological Survey (USGS) has provided a time series of reflectance measures of the Earth surface (Eidenshink and Faundeen, 1994; Loveland and Belward, 1997), intended as a tool for studying long-term changes in vegetation related to global climate changes and anthropic interactions (Eidenshink, 1992). A number of synthetic reflectance indices related to vegetation photosynthetic activity have been subsequently proposed and adopted in vegetation studies, interpretable as the ecologic response of vegetation to the ecosystem changes. NDVI (the normalized ratio of near-infrared and red surface reflectances) is certainly the best known (Tucker, 1979; Goward *et al*., 1991; Ruimy *et al*., 1994; Carlson and Ripley, 1997), and it has been shown to be in strong relationship to the following vegetation properties: a) APAR, amount of absorbed photosinthetically active radiation, b) LAI, leaf area index, c) fractional vegetation cover, d) biomass production, e) NPP, net primary productivity, and f) WSIV, water stress index for vegetation (Cihlar *et al*., 1991; Ricotta *et al*., 1999; Vertovec *et al*., 2001).

According to Holben (1986), single date remote sensing studies can be affected by a series of common problems concerning the goodness of the reflectance measures, such as atmospheric attenuation or cloud contamination of the signal, and surface directional reflectance. Taddei (1997) showed that it is possible to minimize them considering the maximum value assumed by NDVI during a given time period (MVC, Maximum Value Composite method) as a good estimate of the entire period.

For a given region, whose extension is established by the satellite sensor resolution, plotting the NDVI values sequence against time makes it possible to quantitatively analyse the remotely sensed vegetation dynamics. Many techniques have been applied to model vegetation phenological cycle, from simplest parameters like NDVI profiles mean and standard deviation (Ramsey *et al*., 1995), to the onset and peak of greenness and the length of the growing season (Odenweller and Johnson, 1984; Lloyd, 1990; Loveland *et al*., 1991; Reed *et al*., 1994; Running *et al*., 1994). Long NDVI time series have been submitted to classic time series analysis techniques, such as logarithmic and exponential smoothing (Badhwar and Henderson, 1985; Baret and Guyot, 1986), principal component analysis (Townshend *et al*., 1985; Tucker *et al*., 1985; Eastman and Fulk, 1993; Benedetti *et al*., 1994; Anyamba and Eastman, 1996; Hirosawa *et al*., 1996) or Fourier transformation (Menenti *et al*., 1993; Andres *et al*., 1994; Sellers *et al*., 1994; Taddei, 1997).

Recently many remote sensing studies have been carried out using NDVI for studying arid and semi-arid environment. Some of these have focused on identification of indicators for land degradation processes (e.g.

Gamo, 1999; Karnieli, 2000).

The aim of this study, funded by the Italian Ministry of Agricultural and Forestal Politics, is to identify a number of observation areas in southern Italy (Permanent Reflector Sites, PRS), where vegetation response to water availability will be permanently monitored, to detect anomalous meteo-climatic events producing significant decreases in the expected vegetation reflectance response pattern, which could be regarded as potential drought events.

Data

The Italian Ministry of Agriculture and Forestry Policy, which has funded this research, has provided 12 year-long (1989-2000) time series for: a) daily rainfall, b) daily maximum and minimum temperature, c) daily available water content in soil (AW), d) daily real and potential evapo-transpiration, e) daily NDVI. Each time series was georeferenced to a squared cell of 8x8 km; the cells are the basic territorial units of the national agro-meteorological system model (AGRIT, see AA. VV., 1992). The values referred to each cell were derived from the raw data by means of interpolation (rainfall and temperature: from over 300 meteorological stations data), merging (NDVI data, acquired with AVHHR sensor at a resolution of ca 1.1 km, with a theoric maximum of 64 pixel-based data for each cell) or modelling (AW, available water content in soil, and evapo-transpiration). The cells land cover was also considered, and expressed as the fraction of pixels covered by one of three general categories: a) crops, b) grasslands and pastures, c) forests and woodlands.

A GIS database with over 1 million records of bioclimatic data, NDVI maps, and land cover maps, was created with MS Access by Microsoft and managed with the GIS software ArcGIS by ESRI (Mitchell, 1999). It regards 600 cells of 4 macro-areas of Italy endangered by drought events (Fig. 1): a) the state-owned coastal woods of Lazio region; b) The National Park of Gargano, a protected promontory covered by coppices and high trunk forests in Puglia; c) the Macomer area in Sardinia region, a scarcely anthropic impacted zone, mainly occupied by pastures; and d) the agrarian inlands of Sicily, in the Caltanissetta and Enna districts, traditionally used for the extensive mono cultivation of wheat.

The method

In this study a new method to evaluate drought risk is proposed. Considering the drought dynamics in time and space, a drought risk model must be referred to a defined temporal and spatial extent. Data sampling was performed in order to refer the observations to a set of Operational Geographic Units (OGUs, Crovello, 1981;

Feoli and Zuccarello, 1996), with defined area and land cover class. On the basis of the time series analysis, a new concept is introduced, that is the temporal analogous of an OGU. Given a fixed length window, which shifts step-by-step along a time series, it is possible to single out a time unit at each step. If the window length is calculated according to the seasonality of the time series, the set of time units is homogeneous in periodicity, in the sense that the time units are not affected by the seasonal variability of the original time series. Thus the time units can be defined as Operational temporal Units (OtUs). Drought risk is here defined on the basis of a) the geographic location and land cover class of Operational Geographic Units, and b) the bioclimatic pattern of Operational temporal Units. Data processing was aimed to: a) estimate the characteristic time pattern of each variable in each OGU, in terms of its seasonality and cycle components; b) estimate significant deviations from the characteristic patterns; and c) evaluate them in terms of drought risk, on the basis of the bioclimatic history of the OGUs.

A sample of 16 OGU was extracted from the dataset, according with the following criteria: a) selection of at least one cell for each macro-area and for each land cover class; b) selection of the cells whose time series had the maximum degree of completeness, defined as the amount of not missing or invalidated data; c) selection of the cell with the maximum degree of land cover homogeneity, defined as the fraction of its area covered by a given land cover class; d) after at least one cell was selected, definition of the OGU as the area of the cell covered by the given land cover class. The sampling procedure is illustrated in Fig. 2, and the list of the selected OGU is reported in Tab. 1.

For each of the 16 selected OGUs, the time series of the bioclimatic and remote sensed parameters were extracted from the database. To minimize the problems concerned with the goodness of the reflectance measures, the method of Maximum Value Composite (Taddei, 1997), based on a decadal time scaling, was used. The bioclimatic parameters were also rescaled on decadal basis, considering the cumulative decadal rainfall and evapo-transpiration, and the average decadal temperatures and AW. The dataset for the next step of data processing was therefore constituted, for the 16 OGUs, by time series of the 7 bioclimatic variables, each including 432 decadal data.

The periodic components of the time series (i.e. cyclic seasonal or pluriannual variations), and the fitting curves were calculated for all the bioclimatic variables of each OGU by means of the spectral analysis, with the FFT (Fast Fourier Transformation) method. It considers the estimate of a time series values as a multiple regression problem, in which the dependent variable is the time series, and the independent ones are the sinusoidal functions of all the possible frequencies. The analysis aims to decompose the original series into a

sum of sinusoids with different frequencies, to individuate those that are more related with the original data. The general model of multiple regression is:

$$X_t = a_0 + \sum_{i=1}^{n} [a_i \cdot \cos(\lambda t) + b_i \cdot \sin(\lambda t)] \quad \text{(I)}$$

where λ is the frequency of the $i^{\text{-th}}$ sinusoid. The sine (a_i) and cosine (b_i) parameters are the regression coefficients of the respective functions with the data.

Given the spectral analysis results (reported below in the results section), which showed a common time yearly periodicity model for all of the time series, it was possible to introduce a window of 36 decades moving along the time series. Since the length of the time series was equal to 432 decades, shifting the window decade by decade along the time series produced (432-36+1)=397 OtUs (Fig. 3).

To refer to the appropriate OtU the pattern of each variable in the past 36 decades, the definite integral of the curve fitting the 36 punctual data was calculated, according with the following formula:

$$I_x^i = \int_{t=i-35}^{t=i} j_x(t) \quad \text{(II)}$$

where i is the final decade of the OtU, x is the bioclimatic variable time series, and f (t) is the function fitting x, calculated with the multiregressive model (I).

The definite integral here is intended as a measure of the overall magnitude of the variable in the OtU, and it is the only parameter used to express the pattern of the variable during the previous year. In a further paper, its use will be discussed in detail, also taking into account methods estimating the variable fluctuations within the OtU.

For each OGU a table was compiled, reporting integrals calculated for all the OtUs (rows) and all the bioclimatic variables (columns). All the tables were standardized by column, following the formula shown below, to rescale all the variables to the same magnitude.

$$stdI_x^i = \left(I_x^i - m_{I_x}\right) \Big/ s_{I_x}$$

where I_x^i is the defined integral of the variable x referred to the i-th OtU, m_{Ix} and s_{Ix} are the mean and standard deviation of the frequency distribution of Ix.

In the tables, each OtU is described by a row vector, which summarizes the pattern of the bioclimatic variables recorded (NDVI, maximum and minimum temperatures, rainfall, potential and effective evapo-transpiration, and available water content in soil) in the last 36 decades (Tab. 2). Each bioclimatic variable, after standardization, ranges between –3 and +3, with p-level equal to 0.99865. Therefore it is possible to classify the OtUs according with their bioclimatic anomalies, and evaluate the drought risk associated with a particular

bioclimatic anomalous occurrence.

Since the relationship between the bioclimatic variables and the roles they play in giving rise to a drought event are not formally known, an Artificial Neural Network classifier is proposed, to estimate the drought risk associated with each particular combination of bioclimatic values of the OtUs. The training dataset was created *ad hoc*: it was a fictitious table whose row vectors were filled with extreme values, depicting conditions of maximum (Tab. 3), and minimum drought risk (i.e. maximum availability of water resources, see Tab. 4), giving the corresponding drought risk range from –1 to 1. The ANN model used was a three-level Multilayer Perceptron (MLP), being the input layer the defined integral values of the 7 bioclimatic variables in the given OtU, the second layer the processing core, and the third one the drought risk value (Fig. 4) The net was trained with the Back Propagation (BP) algorithm, using the neural network module of Statistica software by StatSoft Inc.

The drought risk index (DRI) was calculated for all the OtUs of the 16 OGUs. A time-dependent graph was produced for each OGU, showing the DRI time pattern, to identify the relative maxima in the drought risk patterns, and evaluate differences between the responses of the OGUs to the bioclimatic patterns. A cluster analysis of the OGUs was performed, based on their drought risk time pattern, to verify if groups of cells with similar responses are grouped according with their geographic location and/or their land cover class.

Results and Discussion

The spectral analysis results showed a common pattern for all the time series of the 16 OGUs; in Fig. 5 the periodograms of the OGU 3175 (Castelporziano area, forests and woodlands land cover class) are reported as a general example.

All the periodograms showed a strong yearly periodicity (P=36 decades). The differences in all the 112 results (16 OGU x 7 bioclimatic time series) can be summarized by three main fundamental patterns: a) dominance of a strong yearly periodicity, with absence of other significant signals (generally observed in temperatures and evapo-transpiration graphs); b) relative maximum at yearly periodicity, with a major peak at the six-monthly one, a typical mediterranean pattern in rainfall time series; c) preponderance of the yearly periodicity, with relative maxima corresponding to short and long period components (AW and NDVI time series). The magnitude of those secondary peaks is generally negligible, except for those long period sinusoids that are relatively strongly related with the original data. Nevertheless the ratio between their period (> 100 decades) and the length of the time series (432 decades) is too small to assess if such peaks recur along the series

(4-5 times) due to an effective long term cyclicity in the bioclimatic variables time pattern, rather than to the presence of some outlayers separated by relatively long time lags in the data. This issue should be separately investigated, studying longer time series.

The integrals indexes of the bioclimatic variables and the drought risk time pattern in the OGU 3175 (Castelporziano area, forests and woodlands land cover class) are reported in Fig. 6, to be compared on the same time scale. Relative maxima of drought risk can be observed in the OtUs ending in summer of ’93, ’95 and 2000, in correspondence of the local minima in rainfall and AW time pattern. Other minor peaks are evident during the 1990, 1991 and 1992, when higher maximum temperatures and evapo-transpiration were observed. The extremely low evapo-transpiration and the very high water availability in soil are dated in 1996, corresponding to the lower drought risk (i.e.: water abundance) values. Comparing the drought risk time pattern with those of the bioclimatic variables, it is important to stress not only the value of the index (drought risk entity), but also its duration, defined as the number of OtUs in which it persist to be anomalously high (drought risk persistence).

The drought index time pattern of some of the OGUs are reported in Fig. 7. A general common model can be observed, characterized by the relative maxima in ’93, ’95 and 2000, as previously described. The differences between the OGUs can be evaluated observing the result of the numerical classification (Fig. 8), based on their drought index time patterns. Four main clusters were formed, corresponding to the main macro-areas where the OGUs are located. The intra-groups homogeneity is significantly high, since the distance between the elements in each cluster is always smaller than 0.1, which, given the distance measure used in the cluster analysis, corresponds to a correlation equal to 0.9, with 396 degree of freedom. (p-level < 0,05). Nevertheless the OGUs form a heterogeneous sample, due to their geographic location, land cover class, and also climate. In fact, although they are macro-climatically classifiable as mediterranean, some local differences exist, and are confirmed by the different entities observed in the OGUs climatic time series (i.e.: rainfall and temperatures). In this respect, the drought risk index proposed (and also the integral indexes for the bioclimatic variables), estimates water availability anomalies (increasing or decreasing) that appear to be almost synchronous in all the sampled macro-areas, independently from their average climatic regimen, vegetation coverage, soil nature.

Conclusions

In the situations described by the data, the integrals indexes introduced have been proved to be effective in detecting anomalous time patterns along the bioclimatic time series.

The proposed drought risk index seemed to be able to summarize the multivariate relationships involved in

the occurrence of drought events, not revealing punctiform (in space and time) water deficits, but identifying macro- areas, as groups of significantly clustered territorial units, where synchronous patterns of bioclimatic anomalies are likely to occur.

The index, ranging between –1 and +1, is informative not only of drought risk entity, but also of its persistence. In any case, it is neither significantly affected by local land use and vegetation, even if expressed in macro-classes, nor by soil characteristics. Thus, it seems to be useful to manage water resources at large scale, rather than predicting local drought risk.

For each of the geographic areas the behaviour of DRI in the 12 years, showed by its time-dependent graph, could be considered as a "finger print for drought" of the area, and its permanent monitoring could be helpful in programming the local water resources management.

References

AA. VV. (1992) Agricoltura e telerilevamento: Statistiche agricole per il 1992. Monografia del Ministero dell'Agricoltura e Foreste (MAF), Roma, Italy

Andres L, Salas WA, Skole D (1994) Fourier analysis of multi-temporal AVHRR data applied to a land cover classification. Int J Remote Sens 15:1115-1121

Anyamba A, Eastman JR (1996) Interannual variability of NDVI over Africa and its relation to El Niño/Southern Oscillation. Int J Remote Sens 17:2533-2548

Badhwar GD, Henderson KE (1985) Application of thematic mapper data to corn and soybean development stage estimation. Remote Sens Environ 17:197-201

Baret F, Guyot G (1986) Suivi de la maturation de couverts deble par radiometrie dans les domaines visible et proche infrarouge. Agronomie 6:509-516

Benedetti R, Rossini P, Taddei R (1994) Vegetation classification in the middle Mediterranean area by satellite data. Int J Remote Sens 15:583-596

Carlson TN, Ripley DA (1997) On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote Sens Environ 62:241-252

Cihlar J, St-Laurent L, Dyer JA (1991) Relation between the normalized difference vegetation index and ecological variables. Remote Sens Environ 35:279-298

Crovello TJ (1981) Quantitative Biogeography: an Overview. Taxon 30:563-575

Eastman JR, Fulk M (1993) Long sequence time series evaluation using standardized principal components. Photogramm Eng Remote Sens 59:991-996

Eidenshink JC (1992) The 1990 conterminous US AVHRR data set. Photogramm Eng Remote Sens 58:809-813

Eidenshink JC, Faundeen JL (1994) The 1 km AVHRR global land data set-first stages in implementation. Int J Remote Sens 15:3443-3462

Feoli E, Zuccarello V (1996) Spatial pattern of ecological processes: the role of similarity in GIS applications for landscape analysis. In: Fisher M, Scholten H, Unwin D (eds) Spatial Analytical Perspectives on GIS. Taylor and Francis, London, pp 175-185

Feoli E, Giacomich P, Mignozzi K, Oztürk M, Scimone M (2003) Monitoring desertification risk with an index integrating climatic and remotely-sensed data: An example from the coastal area of Turkey. Management of Env Qual 14:10-

21

Gamo M (1999) Classification of arid regions by climate and vegetation. J of Arid Land Studies 9:17-26

Gopal S, Woodcock CE, Strahler AH (1999) Fuzzy Neural Network Classification of Global Land Cover from a 1° AVHRR Data Set. Remote Sens Environ 67:230-243

Goward SN, Markhan B, Dye DG, Dulaney W, Yang J (1991) Normalized difference vegetation index measurements from the Advanced Very High Resolution Radiometer. Remote Sens Environ 35:257-277

Hayes MJ (2003) Drought Indices. NDMC University of Nebraska, Lincoln, http://www.drought.unl.edu/whatis/Indices.pdf

Henderson CE (2000) Predicting Aflatoxin Contamination in Peanuts: A Genetic Algorithm/Neural Network Approach. Appl Intell 12:183-192

Hirosawa Y, Marsh SE, Kliman DH (1996) Application of standardized principal component analysis to land-cover characterization using multitemporal AVHRR data. Remote Sens Environ 58:267-281

Holben BN (1986) Characteristics of maximum-value composite images from temporal AVHRR data. Int J Remote Sens 7:1417-1434

Hoptroff R (1993) The principles and practice of times series forecasting and business modelling using neural nets. Neural Comput Appl 1:59-66

Karnieli A (2000) Drought monitoring in the Negev desert using NOAA/AVHRR imagery. Proceeding of Agroenvironment 2000, II International symposium on New Technologies for Environment Monitoring and Agro-Applications, Tekirdag, Turkey

Keetch JJ, Byram GM (1968) A Drought Index for Forest Fire Control. Research Paper 38, US Dept of Agriculture, Washington

Kuo C, Reitsch A (1995) Neural networks vs conventional methods of forecasting. J Business Forecast Meth Syst 14:17-22

Kwak NK, Changwon L (1997) Neural Network Application to Classification of Health Status of HIV/AIDS Patients. J Med Syst 21:87-97

Laepes A, Farben R (1987) Nonlinear signal processing using neural networks: prediction and system modelling. Tech Rep, Los Alamos National Laboratory, Los Alamos, New Mexico

Lloyd D (1990) A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. Int J Remote Sens 11:2269-2279

Loveland TR, Belward AS (1997) The IGBP-DIS 1 km land cover data set, DISCover: first results. Int J Remote Sens 18:3289-3295

Loveland TR, Merchant JW, Ohlen DO, Brown JF (1991) Development of a land-cover characteristics database for the conterminous US. Photogramm Eng Remote Sens 57:1453-1643

McKee TB, Doesken NJ, Kleist J (1993) The relationship of drought frequency and duration to time scales. Proceedings of the VIII Conference on Applied Climatology, Boston. Am Met Soc:179-184

Menenti M, Azzali S, Verhoef W, Van Swol R (1993) Mapping agro-ecological zones and time lag in vegetation growth by means of Fourier analysis of time series of NDVI images. Adv in Space Res 13:233-237

Mitchell A (ed) (1999) The ESRI guide to GIS analysis. ESRI Press, Redlands, California

Moshiri S (1999) Static, Dynamic, and Hybrid Neural Networks in Forecasting Inflation. Comput Econ 14:219-235

Odenweller JB, Johnson KI (1984) Crop identification using Landsat temporal-spectral profiles. Remote Sens Environ 14:39-54

Palmer WC (1965) Meteorological Drought. Research Paper 45, US Dept of Commerce, Washington

Ramsey RD, Falconer A, Jensen JR (1995) The relationship between NOAA-AVHRR NDVI and Ecoregions in Utah. Remote Sens Environ 53:188-198

Reed BC, Brown JF, Van der Zee D, Loveland TR, Merchant JW, Ohlen DO (1994) Measuring phenological variability from satellite imagery. J Veg Sci 5:703-714

Ricotta C, Avena G, De Palma A (1999) Mapping and monitoring net primary productivity with AVHRR NDVI time-series:

statistical equivalence of cumulative vegetation indices. J Photogramm Remote Sens 54:325-331

Ruimy A, Saugier B, Dedieu G (1994) Methodology for the estimation of terrestrial net primary production from remotely sensed data. J Geophys Res 99:5263-5283

Running SW, Loveland TR, Pierce LL (1994) A vegetation classification logic based on remote sensing for use in global biogeochemical models. Ambio 23:77-81

Sellers PJ, Tucker CJ, Collatz GJ, Los SO, Justice CO, Dazlich DA, Randall DA (1994) A global 18 by 18 NDVI data set for climate studies: Part 2. The generation of global fields of terrestrial biophysical parameters from the NDVI. Int J Remote Sens 15:3519-3545

Svoboda M, LeComte D, Hayes M, Heim R, Gleason K, Angel J, Rippry B, Tinker R, Palecki M, Stooksbury D, Miskus D, Stephens S (2002). The drought monitor. Bull Am Meteor Soc 83:1181-1190 DOI 10.1175/1520-0477(2002)083

Taddei R (1997) Maximum value interpolated MVI: a maximum value composite method improvement in vegetation index profile analysis. Int J Remote Sens 18:2365-2370

Tadesse T, Wilhite DA, Harms SK, Hayes MJ, Goddard S (2004) Drought Monitoring Using Data Mining Techniques: A Case Study for Nebraska, USA. Natural Hazards 33:137-159

Thornthwaite CW (1948) An approach towards a rational classification of climate. Geographical Review 38:55-94

Townshend JRG, Goff TE, Tucker CJ (1985) Multitemporal dimensionality of images of normalized difference vegetation index at continental scales. IEEE Trans Geosci Remote Sens 23:888-895

Tucker CJ (1979) Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens Environ 8:127-150

Tucker CJ, Townshend JRG, Goff TE (1985) African landcover classification using satellite data. Science 227:369-375

Vertovec M, Saçkali S, Oztürk M, Salleo S, Giacomich P, Feoli E, Nardini A (2001) Diagnosing plant water status as a tool for quantifying water stress on a regional basis in Mediterranean drylands. Ann For Sci 58:113-125

Wasserman P (ed) (1989) Neural Computing: Theory and Practice. Van Nostrand Reinhold, New York

White H (ed) (1992) Artificial Neural Networks: Approximation and Learning Theory. Blackwell, Oxford

Wilhite DA (1992) Drought. Enc of Earth System Science 2:81-92

Wilhite DA (2000) Reducing societal vulnerability to drought. In: Wilhite DA (ed) Drought: A Global Assessment. Routledge, London, pp 295-298

Wu H, Wilhite DA (2004) An Operational Agricultural Drought Risk Assessment Model for Nebraska, USA. Natural Hazards 33: 1-21

Tab. 1. List of the OGUs selected for each macro-area. Homogeneity (H) is calculated as the percentage of the cell area (64 km^2) covered by the land cover class; completeness (C) is calculated as the percentage of decades in the time series for which a valid data was recorded.

OGU	Region	Land cover class	Area (km^2)	Cell code	H	Decades count	C
1	Sicily	Forests/woodlands	44	5074	68.8%	379	87.7%
2	Sicily	Grasslands/pastures	42	5069	65.6%	381	88.2%
3	Sicily	Crops	64	4994	100.0%	380	88.0%
4	Sicily	Crops	64	4986	100.0%	384	88.9%
5	Sardinia	Forests/woodlands	64	4320	100.0%	380	88.0%
6	Sardinia	Forests/woodlands	62	4271	96.9%	380	88.0%
7	Sardinia	Grasslands/pastures	19	3993	29.7%	377	87.3%
8	Sardinia	Grasslands/pastures	15	4291	23.4%	373	86.3%
9	Sardinia	Crops	58	4267	90.6%	382	88.4%
10	Sardinia	Crops	49	4172	76.6%	378	87.5%
11	Lazio	Forests/woodlands	55	3175	85.9%	380	88.0%
12	Lazio	Crops	41	3133	64.1%	377	87.3%
13	Puglia	Forests/woodlands	57	3085	89.1%	363	84.0%
14	Puglia	Forests/woodlands	58	3086	90.6%	371	85.9%
15	Puglia	Crops	54	3169	84.4%	375	86.8%
16	Puglia	Grasslands/pastures	17	3168	26.6%	374	86.6%

Tab. 2 – Table of the standardized defined integrals for cell 3175 in the Castelporziano area, forest/woodland fraction. For each OtU (row), the starting and ending decades, and the standardized defined integrals for NDVI, T max, T min, rainfall, potential and real evapo-transpiration and available water content in soil are shown.

OTU	start		end		$stdI_{NDVI}$	$stdI_{Tmax}$	$stdI_{Tmin}$	$stdI_{R}$	$stdI_{ETP}$	$stdI_{ETR}$	$stdI_{AW}$
	year	decade	year	decade							
1	1989	1	1989	36	-0.2392	1.1045	-1.6105	-0.4669	-0.3373	1.1259	0.5591
…	…	…	…	…	…	…	…	…	…	…	…
395	1999	35	2000	34	-1.6719	-0.3013	1.5834	-0.7305	-0.5040	0.3123	-0.5184
396	1999	36	2000	35	-1.8583	-0.2106	1.6627	-1.0038	-0.4834	0.3279	-0.1437
397	2000	1	2000	36	-2.0989	-0.0492	1.7866	-0.7908	-0.5235	0.3131	0.2009

Tab. 3 – Example of fictitious row vector used to train the ANN classifier. It describes a situation of maximum drought risk: in the hypothetic past 36 decades temperatures and evapo-transpiration have been anomalously high, rainfall and available water content in soil have been anomalously low, and vegetation response to this ecological conditions (NDVI) have been an anomalous decrease of the photosynthetic activity.

$stdI_{NDVI}$	$stdI_{Tmax}$	$stdI_{Tmin}$	$stdI_{R}$	$stdI_{ETP}$	$stdI_{ETR}$	$stdI_{AW}$	DROUGHT RISK
-3	3	3	-3	3	3	-3	**1**

Tab. 4 – Example of fictitious row vector used to train the ANN model. It describes a situation of anomalously high water availability: in the hypothetic past 36 decades temperatures and evapo-transpiration have been low, rainfall and available water content in soil have been anomalously high, and vegetation have increased the photosynthetic activity.

$stdI_{NDVI}$	$stdI_{Tmax}$	$stdI_{Tmin}$	$stdI_{R}$	$stdI_{ETP}$	$stdI_{ETR}$	$stdI_{AW}$	DROUGHT RISK
3	-3	-3	3	-3	-3	3	**-1**

Figures

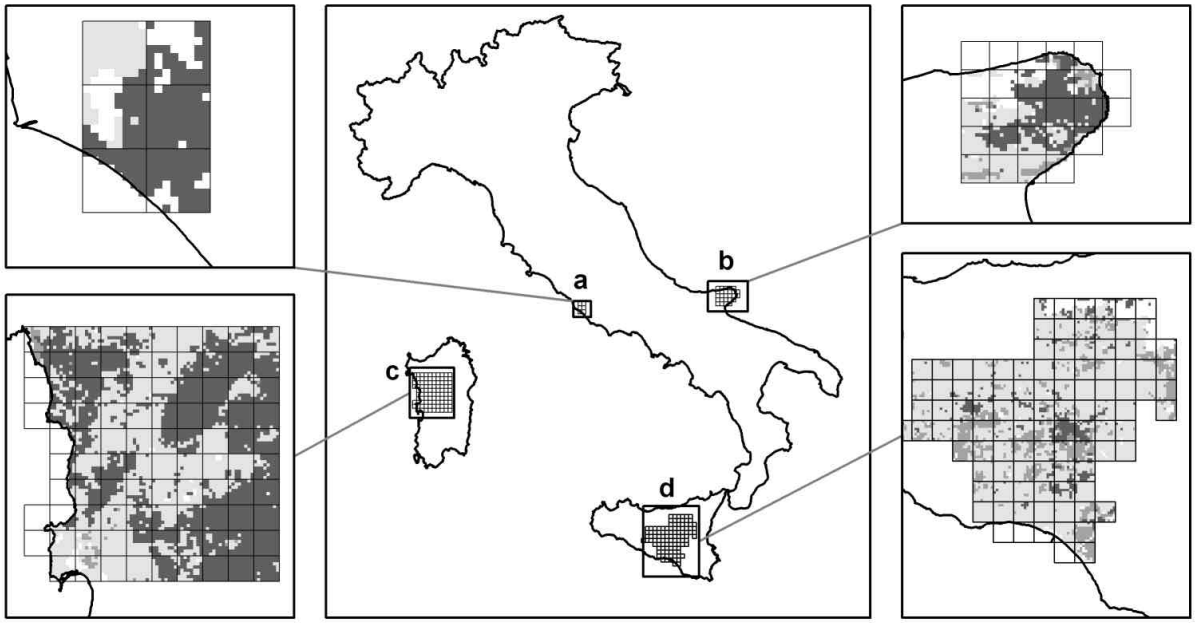

Fig. 1. – 8x8 km squared cells location in the 4 Italian regions endangered by drought: a) Lazio, b) Puglia, c) Sardinia, d) Sicily. Main land cover classes are in dark (forests and woodlands), medium (grasslands and pastures) and light grey (crops); areas not covered by any of the former classes are in white.

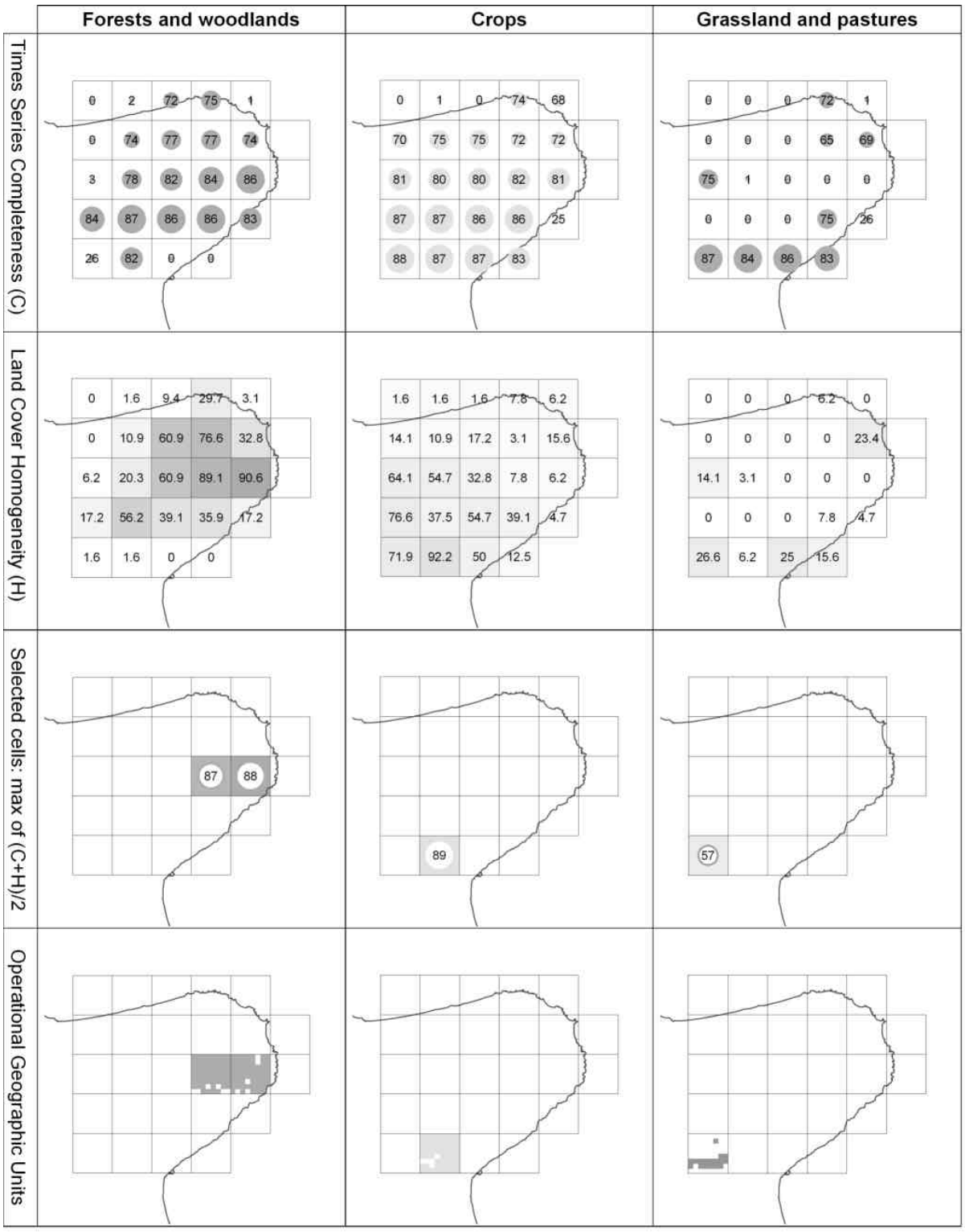

Fig. 2 – OGUs selection in the macro-area of Gargano. For each land cover class (columns), rows show time series completeness of the cells (C, the percent of decades with a valid recorded data); their land cover homogeneity (H, the percent of the cell area covered by the given class), the cells with higher values of C and H, and the 4 OGUs selected for the macro-area.

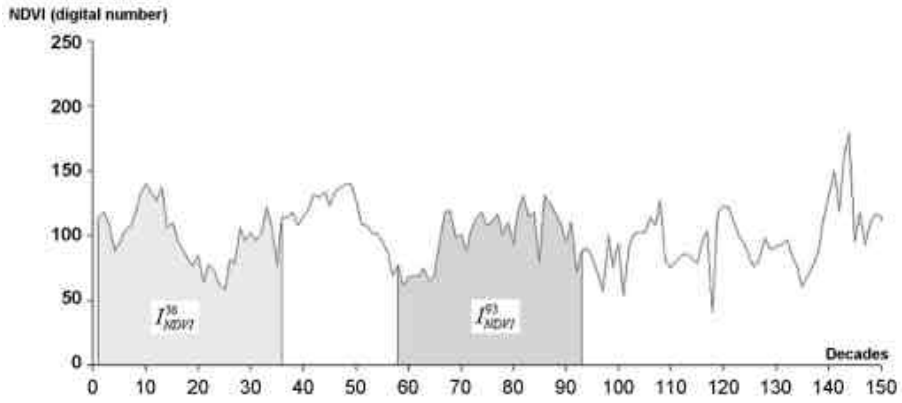

Fig. 3 – Operational temporal Unit (OtU). Drought risk is referred to windows of 36 decades. Here two OtU are shown, for decade 36 and 93. The pattern of bioclimatic variables for each OtU is expressed by the defined integral of the curve fitting the variable profile (see text). Data are from the OGU 3175 (Castelporziano area, forest/woodland fraction, NDVI).

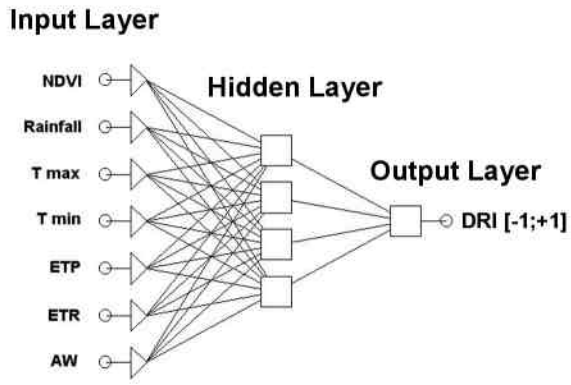

Fig. 4 –ANN model (Three Layer Perceptron) used to estimate the drought risk (see text).

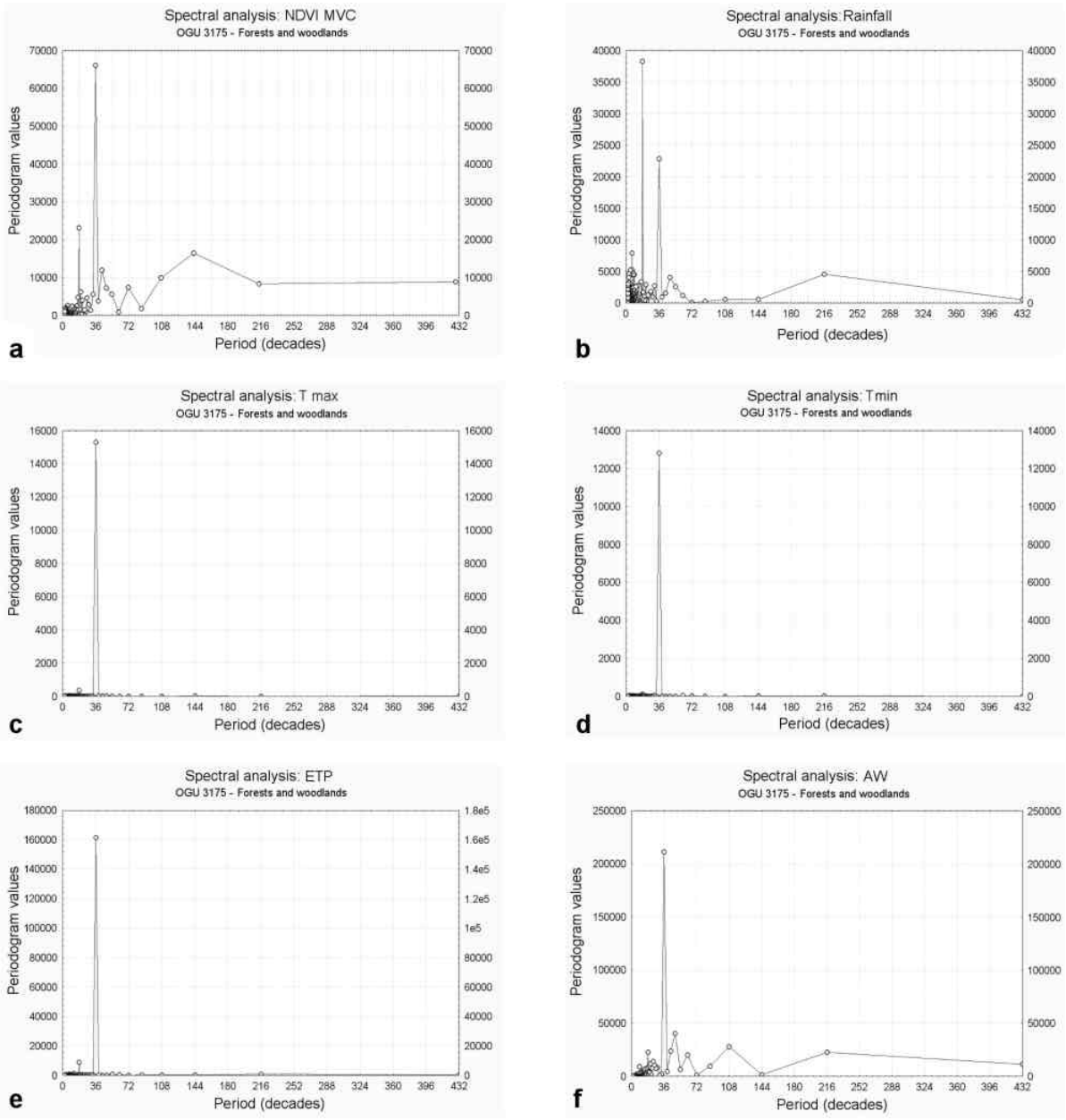

Fig. 5 – Spectral analysis results, OGU 3175 (Castelporziano area, forest/woodland fraction). On the abscissa the period of sinusoidal functions (number of decades). On the ordinate values of the periodogram, calculated as $Pi = a_i^2 \cdot b_i^2 \cdot n/2$, (a, b, and n are as defined in the equation I). P_i measures the amount of periodicity in the original time series explained by the i-th sinusoid.

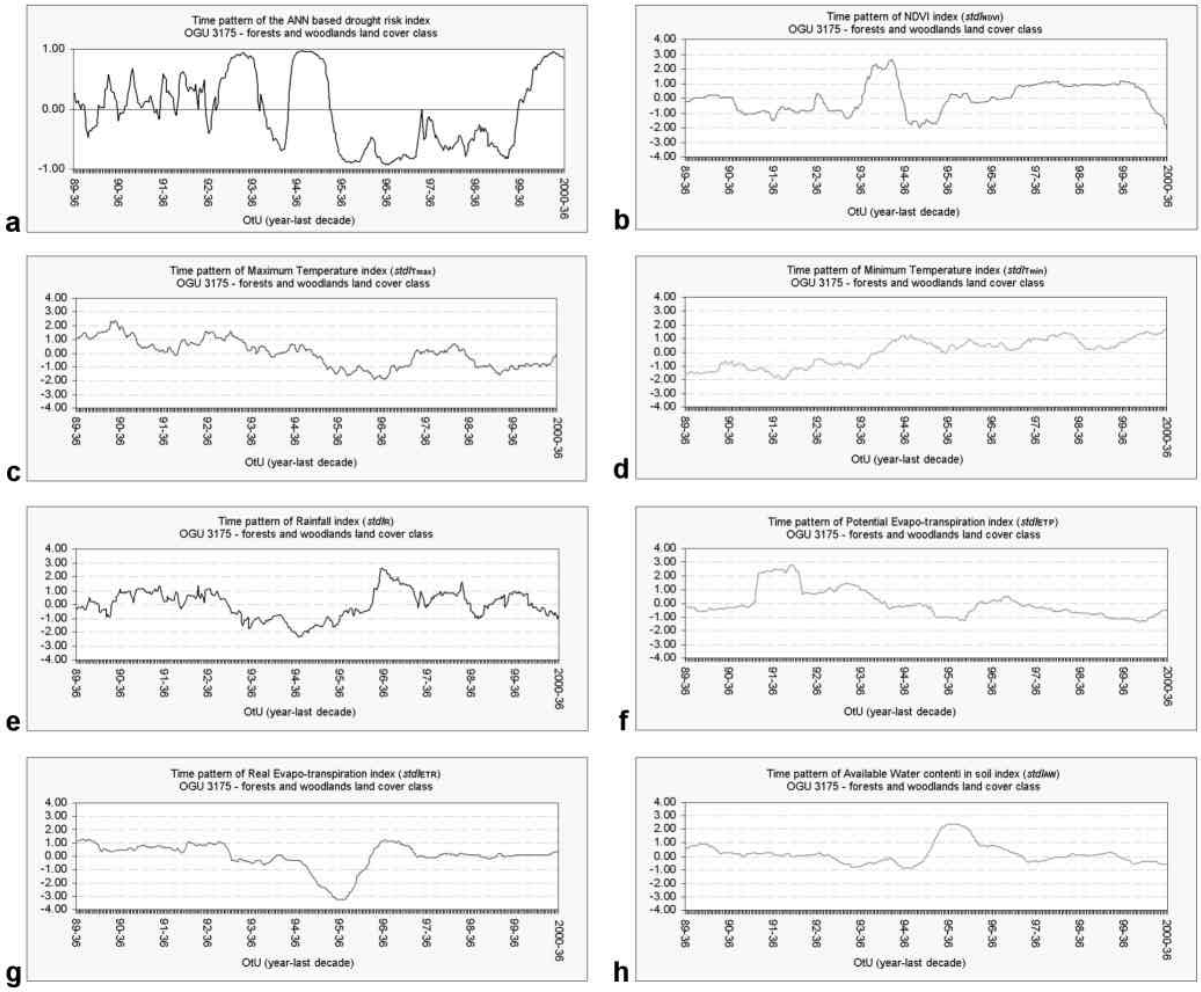

Fig. 6 – Time pattern of DRI (A) and of the 7 bioclimatic variables integral indexes (B-H), OGU 3175 (Castelporziano area, forests and woodlands).

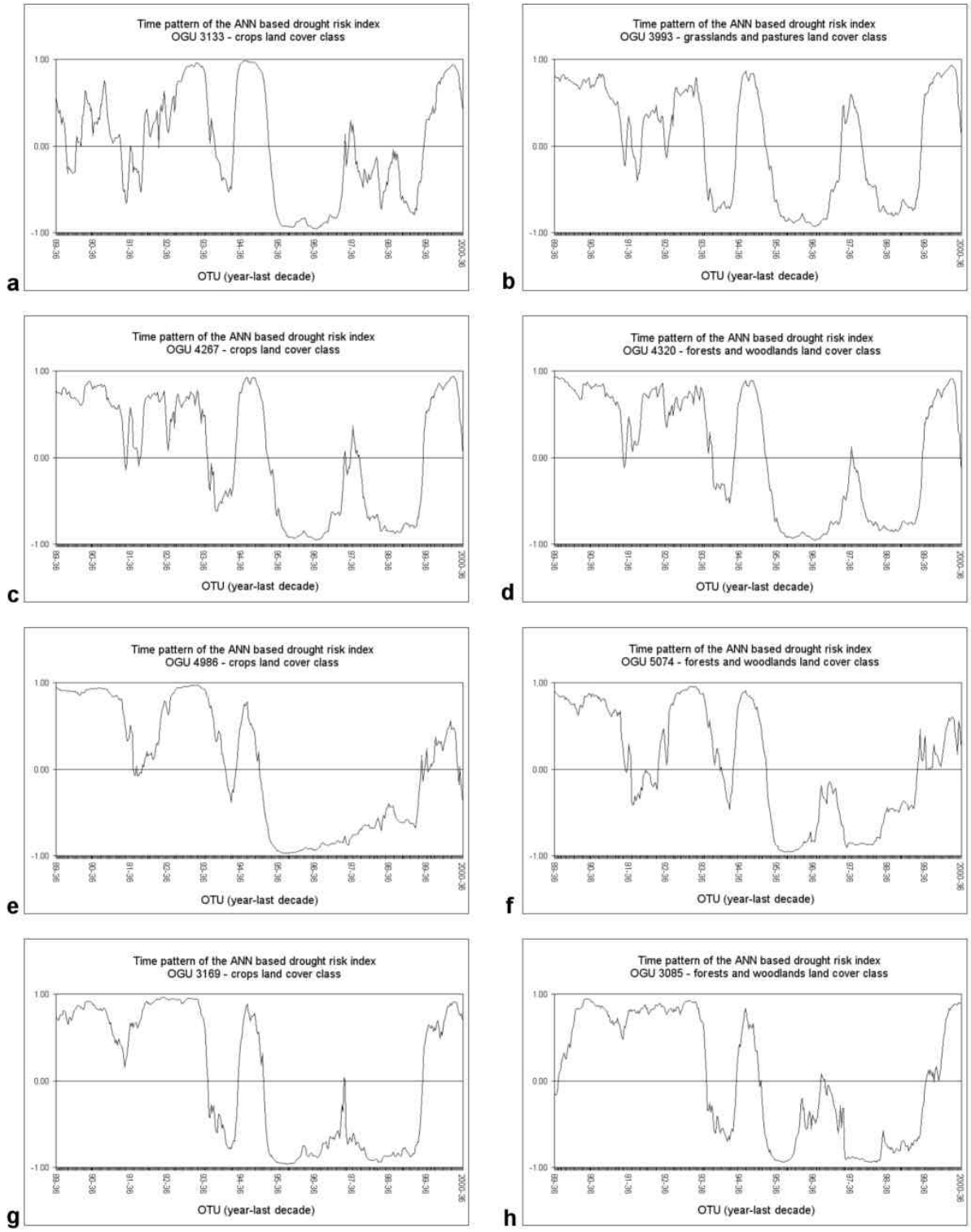

Fig. 7 –Drought risk index time pattern of 8 OGUs. The OGU are named with the same codes as in Tab. 3.

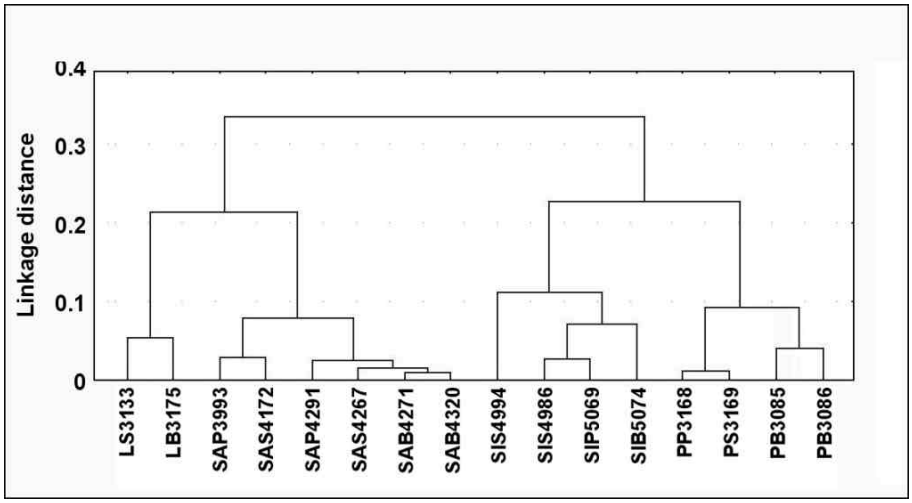

Fig. 8 – Cluster analysis of the 16 OGUs. The OGUs are named by an alphanumeric code built concatenating: a) the macro-area (L = Lazio, P = Puglia, SA = Sardinia, SI = Sicily); b) the land cover class (B = forests and woodland, P = grasslands and pastures, S = crops); c) the database code of the cell.