Spatiotemporal Drought Variability and Classification over Greece based on Remotely Sensed Indices

Nicolas R. Dalezios, N. V. Spyropoulos, A. Blanta and E. Kanellou

Laboratory of Agrometeorology
Faculty of Agricultural Sciences
University of Thessaly

Bratislava, 3-6 May 2011

OUTLINE

- Data Set
- Methodology
  - RDI
  - PCA - Clustering
- Results
  - Tables
  - Images
- Discussion
- Conclusions

Data Set

Satellite Data

- Monthly Brightness Temperature (BT) from channels 4 and 5 of NOAA/ AVHRR satellite of 21years (1981-2001), 8x8 km spatial resolution.
- Monthly NDVI images for the same time period and pixel size.
- Monthly air temperature extracted by LST images

Ground measurements

- Daily precipitation of Greece water district in 50 x 50 km grid size.

Methodology

- *Reconnaissance Drought Index (RDI)*

  Estimate Meteorological Drought conditions based on hydro-meteorological parameters.

- *PCA and Clustering*

Reconnaissance Drought Index

The RDI_{st} is calculated by the equation:

$$RDI_{st}(k) = \frac{y_k - \bar{y}_k}{\hat{\sigma}_k}$$

y_k is the $\ln a_k$, y_k (upper line) is its arithmetic mean and σ_k is its standard deviation

a_k = the initial value for the index (For October $a_k = 1$) and is calculated by:

$$a_k = \frac{\sum_{j=1}^{j=k} p_j}{\sum_{j=1}^{j=k} PET_J}$$

P_j and PET_j are the precipitation and potential evapotranspiration respectively of the j-th month of the hydrological year.

RDI Drought classes

Drought Categories	RDI Values
Extremely Wet	>2.00
Very Wet	1.50 to 1.99
Moderately Wet	1.00 to 1.49
Near Normal	-0.99 to 0.99
Moderately Dry	-1.00 to -1.49
Severely Dry	-1.50 to -1.99
Extremely Dry	<-2.00

RDI methodology

- Land Surface Temperature (LST) calculation from channels 4 and 5 of satellite.
- Air temperature extraction from LST images using air temperature data from meteorological stations.
- Estimation of potential evapotraspiration (ET_p) with Blaney-Criddle method.
- Combination of precipitation maps derived from ground measurements with ET_p maps for RDI extraction.

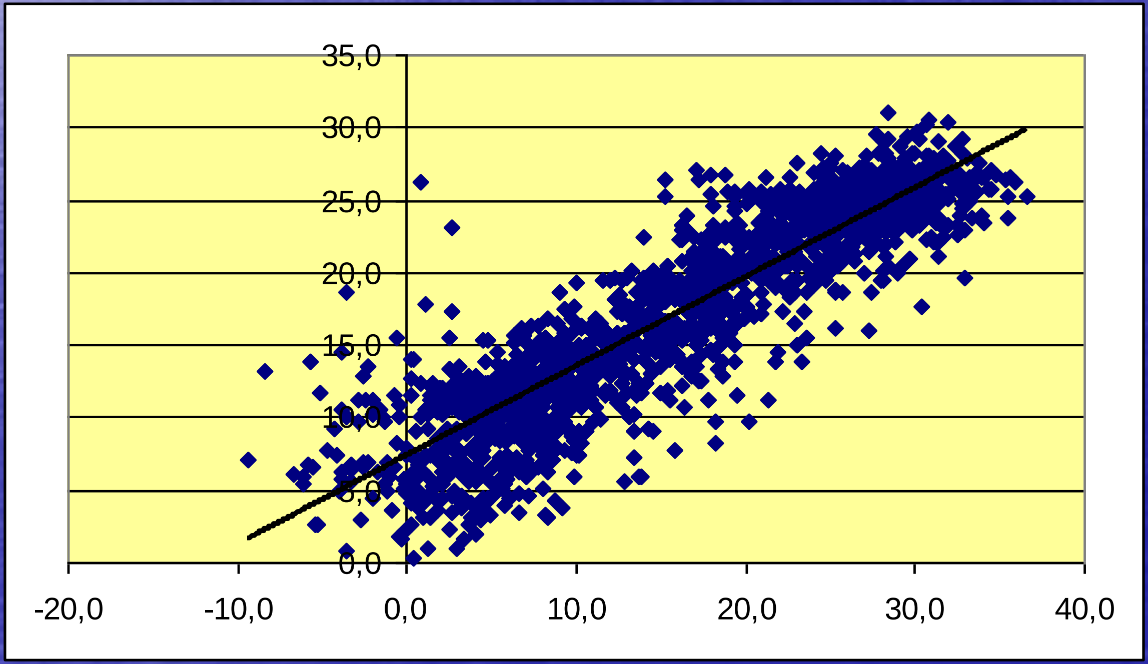
LST

LST extraction for the whole timeseries (1981-2001) in 8 x 8 km pixel size

Air Temperature estimation

- Empirical relationship between LST and air temperature (T_{air}) ($R^2 \approx 0.82$):

$$T_{air} = 0{,}6143 * LST + 7{,}3674$$



ET_p Blaney- Criddle

Estimation of ET_p by the Blaney-Criddle method:

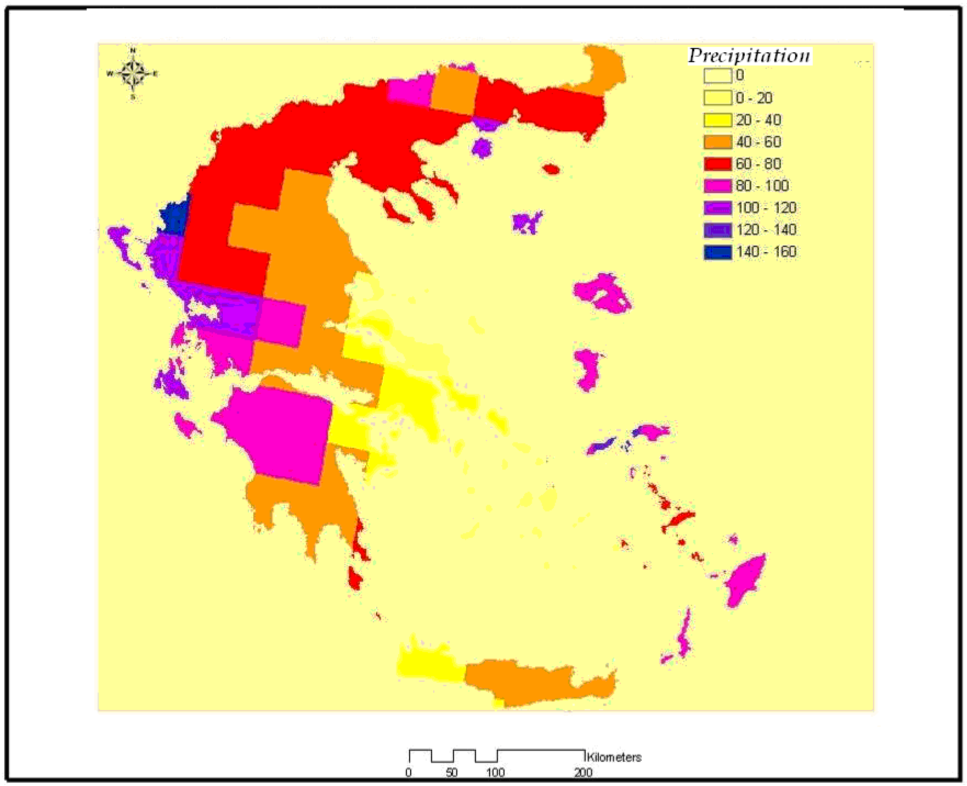
$$ET_p = k * [0.46T+8.16]*p$$

ET_p is the monthly potential evapotranspiration in mm, k is the monthly crop coefficient, T is the mean monthly air temperature (^{o} C) and p is the percentage of day hours

- The Crop coefficients k are calculated for each different vegetation type of the study area based on Corine Hellas 2000 and for each month of the year.
- Extract day hours percentages (p) maps for every month for the region middle Latitude (39^o).
- Both k and p maps are extracted in GIS environment (ArcMap 9.1).

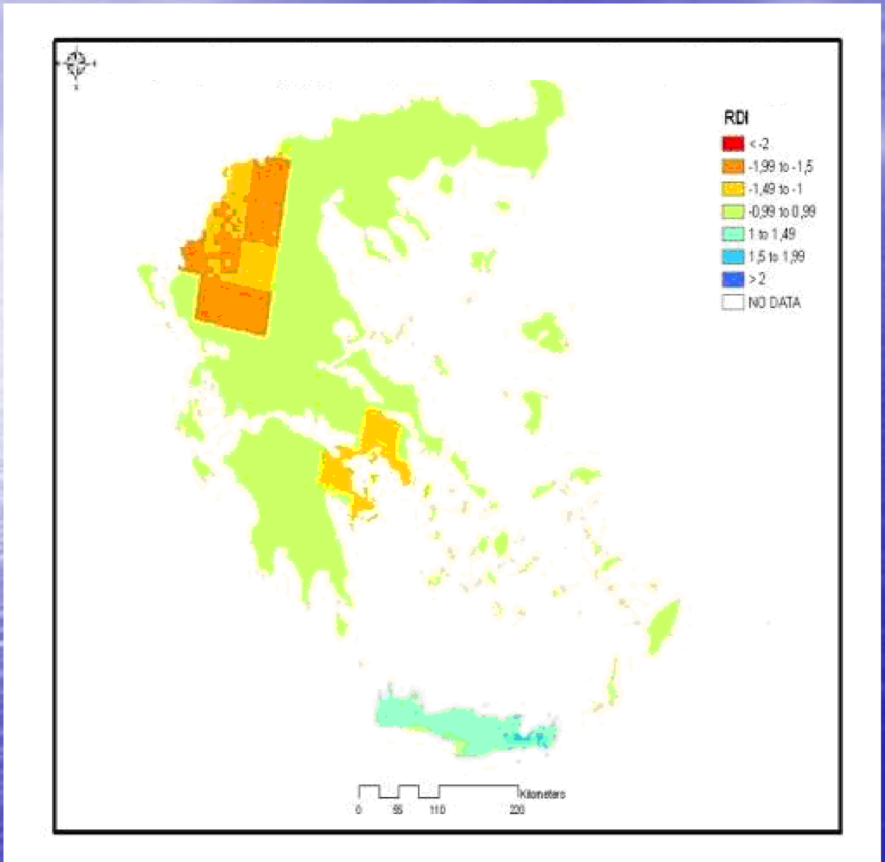
Precipitation maps

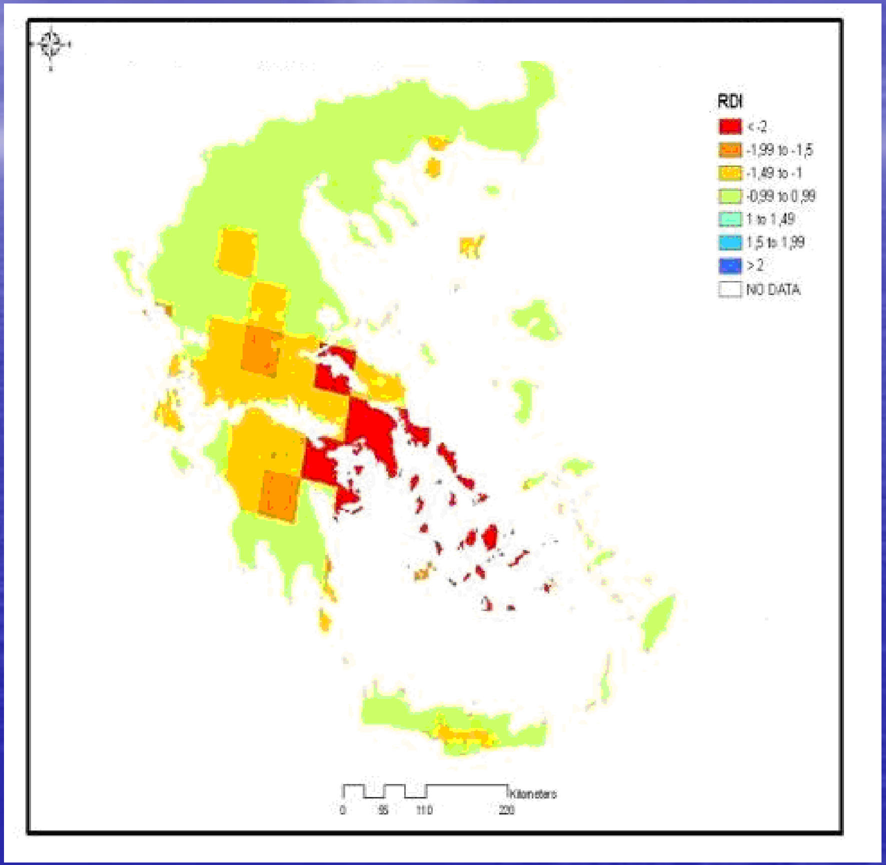
Extraction of monthly rainfall maps based on conventional daily data in 50 x 50 km grid size from 1981 to 2001.



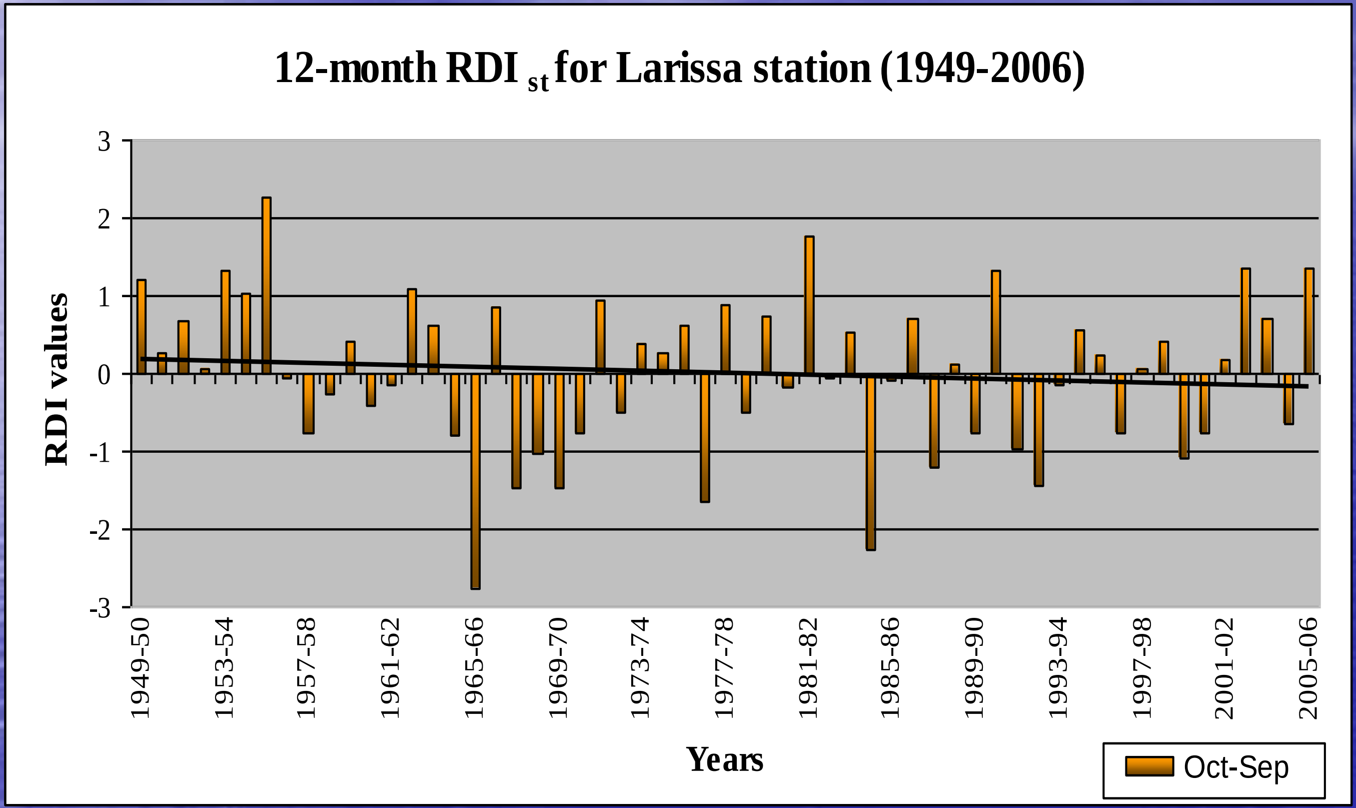
Precipitation map for April 2001

RDI maps





Typical annual RDI Index for the hydrological years 1996-1997 (left) and 1999-2000 (right).



12-month RDI st for Larissa station (1949-2006)
RDI values
3
2
1
0
-1
-2
-3
1949-50
1953-54
1957-58
1961-62
1965-66
1969-70
1973-74
1977-78
1981-82
1985-86
1989-90
1993-94
1997-98
2001-02
2005-06
Years
Oct-Sep

PCA and Clustering

A linear transformation, the new measurement axes are linear combinations of original measurement axes. Using the eigenvectors of the covariance matrix (resulted from PCA) of a data set as new measurement axes for that data set has two major effects.

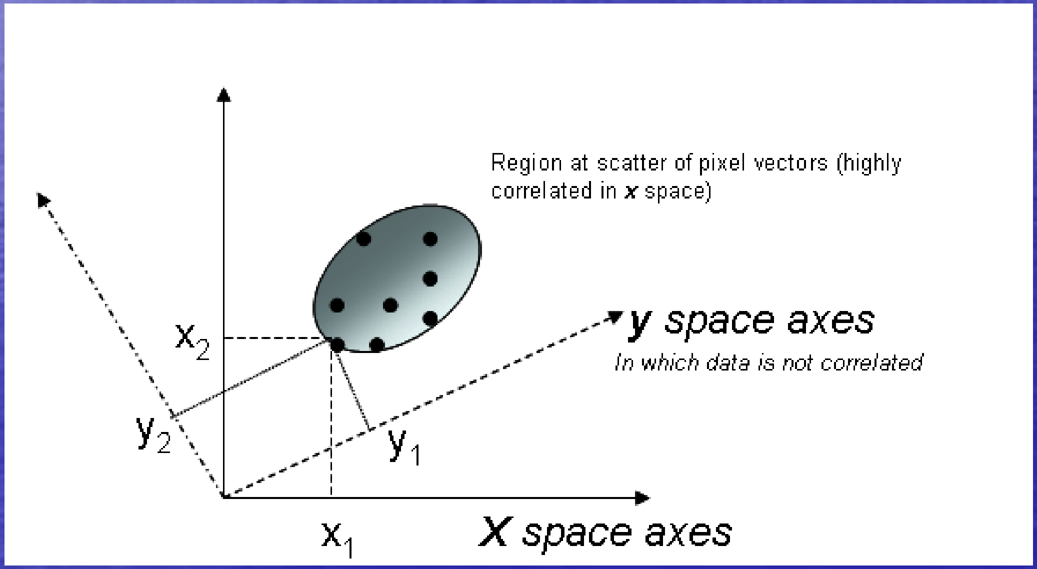


Illustration of a modified coordinate system in which the pixel vectors have uncorrelated components

- First, the new channels are orthogonal with respect to each other, which is not the case with most raw image data since channels are usually highly correlated (particularly the RDI data of our study).
- Second, the variance (information plus noise) implicit in the original channels are “packed” in the new channels such that the eigenchannel (vector) with the highest eigenvalue (eigenchannel 1) typically contains considerably more variance than the second eigenchannel (Richards, 1986).
- The same comparison holds for eigenchannel i and eigenchannel i+1. It has to be mentioned that a result of the principal component transformation is that the new midpoint for each eigenchannel is at 0, with approximately half the new data being negative and half positive. A separate, new midpoint can be specified for each selected eigenchannel (Ingebritsen and Lyon, 1985). In our case the PCA of the 19 RDI annual channels produced the results shown in the following tables.

The eigenchannels 1, 2, 3, 4, 5 and 6 packed for the 92% of the variance from the 19 input channels. The `information' contained in the variance accounts for class discriminability (or lack of it).

Eigenchannel	Eigenvalue	Deviation	Variance (%)
1	25,0778	5,0078	71,46%
2	3,4257	1,8509	9,76%
3	1,8011	1,3421	5,13%
4	1,1064	1,0518	3,15%
5	0,7666	0,8755	2,18%
6	0,5786	0,7606	1,65%
7	0,4694	0,6851	1,34%
8	0,4	0,6325	1,14%
9	0,3314	0,5757	0,94%
10	0,2403	0,4902	0,68%
11	0,1931	0,4395	0,55%
12	0,158	0,3975	0,45%
13	0,1278	0,3574	0,36%
14	0,1157	0,3402	0,33%
15	0,085	0,2915	0,24%
16	0,0754	0,2746	0,21%
17	0,0595	0,2439	0,17%
18	0,0492	0,2219	0,14%
19	0,031	0,1761	0,09%

Covariance matrix of the 5 first eigenchannels of the 19 input annual RDI indices

L/I	Years	PC1	PC2	PC3	PC4	PC5
1	1982-83	0,21719	-0,21139	-0,22244	-0,18778	-0,11119
2	1983-84	0,19979	-0,22228	0,07336	-0,07204	0,11776
3	1984-85	0,23201	-0,44866	0,71884	0,16171	-0,13312
4	1985-86	0,21595	-0,33443	-0,20252	0,11995	0,17003
5	1986-87	0,21771	-0,25736	-0,28833	0,43902	-0,04028
6	1987-88	0,2147	-0,12307	-0,12829	0,11744	-0,05293
7	1988-89	0,23724	-0,04329	-0,17156	0,01515	-0,38073
8	1989-90	0,26058	-0,13881	-0,15949	-0,34753	-0,42039
9	1990-91	0,20961	-0,24184	-0,19323	-0,25575	0,58459
10	1991-92	0,24226	-0,01305	0,21853	-0,1029	0,13123
11	1992-93	0,25576	0,14511	0,19944	-0,30419	0,18401
12	1993-94	0,22658	0,1591	-0,00192	-0,05055	-0,13086
13	1994-95	0,22157	0,23894	0,0869	-0,02185	0,1666
14	1995-96	0,2204	0,20834	-0,02509	0,25025	0,31172
15	1996-97	0,23905	0,23797	0,25419	0,18274	-0,05605
16	1997-98	0,23325	0,26682	0,05564	-0,35986	-0,13429
17	1998-99	0,22002	0,24548	-0,07565	0,1331	0,14908
18	1999-00	0,2468	0,17241	-0,15359	-0,05289	-0,08453
19	2000-01	0,23843	0,25867	-0,05343	0,41701	-0,1249

Clustering

- After the first orthogonal transformation based on PCA analysis the first 6 resulted components are also orthogonally transformed through clustering in order to be categorized to meaningful classes.
- The clustering method selected is the minimum distance (MINDIS) algorithm based on the following K-Means objective function:

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2$$

where $\left\| x_i^{(j)} - c_j \right\|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster centre c_j, is an indicator of the distance of the *n* data points from their respective cluster centres.

K-Means is also based on the minimum Euclidian distance classifier defined by the following equation:

*Gi(X) = (X-Ui)T * (X-Ui)*

*= SUM[(xj-uj)**2] for j = 1 to d.*

Gi(X)	is the result for class i on pixel X
T	indicates transposition of the elements in brackets
D	is the number of channels in the classification
X=(x1,...,xd)	is the (d by 1) pixel vector of grey-levels
Ui=(u1,...,ud)	is the (d by 1) mean vector for class I
J	is the subscript of jth element of a vector
SUM[]	is the total of elements inside brackets

- The distances between the pixel to be classified and each class centre are compared. The pixel is assigned to the class whose centre is the closest to the pixel. If for all i not equal j, $Gj(X) < Gi(X)$, then X is classified as j (Hodgson, 1988). The result of the classification is a theme map directed to a specified database image channel
- Three clustering sessions/scenarios were tested using the 2nd, 3rd, 4th, 5th and 6th components produced by PCA. The first session scheduled with 3 classes the second with 4 and the third with 5 respectively. Tables 4, 5 and 6 summarize the statistics generated from these sessions.

Clustering using K-Means, Number of Clusters: 3

Cluster representation	Pixels	Mean Position	Std Dev
(1) (green)	74844	1.93380 2.18327 0.00458 -0.17735 -0.48133	1.23748 0.68340 1.65674 0.75501 0.50031
(2) (blue)	293674	-1.45711 0.15749 0.00487 0.17670 0.04756	0.93687 0.58671 0.72908 0.84415 0.64409
(3) (yellow)	159663	1.77362 -1.31312 -0.01110 -0.24187 0.13816	0.44305 1.04966 1.18127 0.90957 0.94959
total	528181		

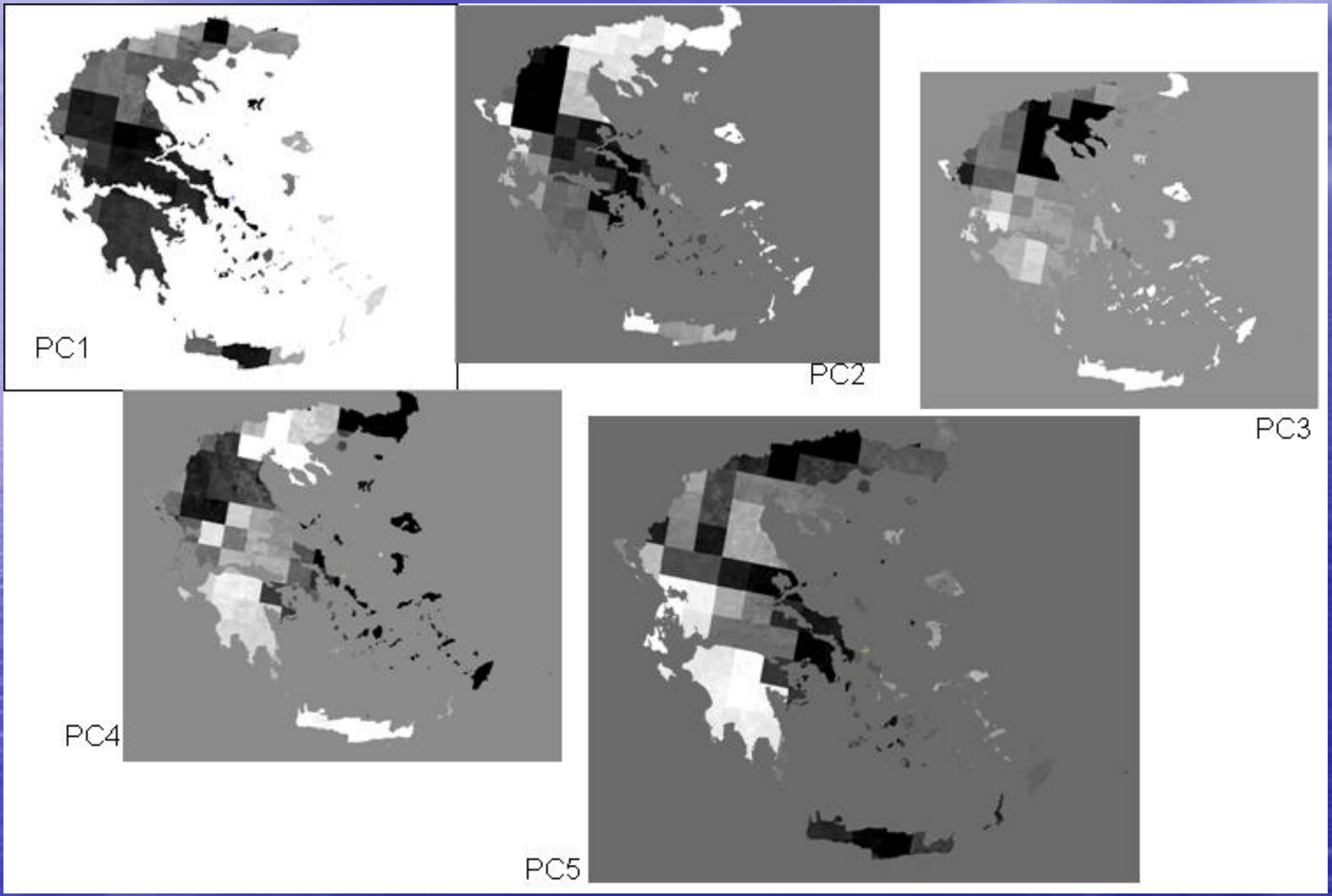
Clustering using K-Means, Number of Clusters: 4

Cluster representation	Pixels	Mean Position	Std Dev
(1) (green)	148586	-0.29488 0.84105 0.81213 0.71149 0.09346	0.61744 0.83436 0.58597 0.86006 0.46404
(2) (blue)	174320	-2.09829 -0.08697 -0.43216 -0.37444 -0.07003	0.57613 0.53772 0.58964 0.50373 0.74620
(3) (yellow)	92446	2.40555 1.07933 -1.09431 -0.22128 0.15103	0.72515 1.20064 1.21752 0.87315 0.98652
(4) (orange)	112829	1.65920 -1.85757 0.49479 -0.17715 -0.13863	0.42380 0.73381 0.76952 0.80181 0.84272
Total	528181		

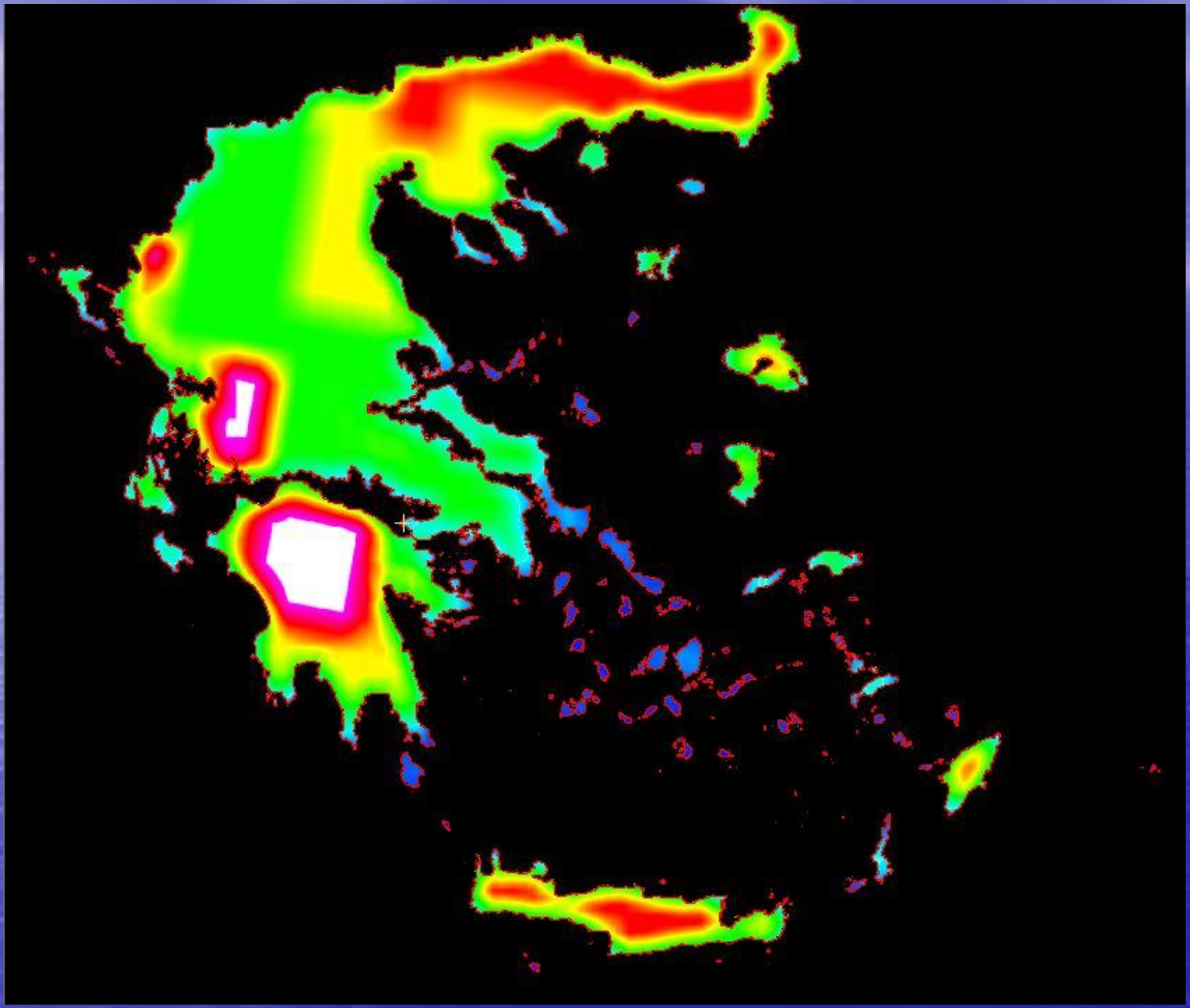
Clustering using K-Means, Number of Clusters: 5

Cluster representation	Pixels	Mean Position	Std Dev
(1) (green)	16995	-0.67833 1.10584 -1.00775 -0.36391 -1.56960	0.69443 0.92025 0.83688 0.47800 0.71886
(2) (blue)	174981	-2.13519 -0.11323 -0.30957 -0.31793 0.06497	0.48922 0.37351 0.57103 0.49690 0.63969
(3) (yellow)	134515	1.12885 -1.40127 0.41680 0.51331 -0.27162	0.87375 1.18055 0.66005 0.65848 0.44462
(4) (orange)	127582	2.07536 1.14424 -0.24135 -0.64019 0.26072	0.92346 1.32112 1.65165 0.80808 0.98020
(5) (pink)	74108	-0.42478 0.58734 0.62099 1.00455 0.25075	0.32742 0.55217 0.45512 0.67445 0.37038
total	528181		

Classification Results



Principal Component Analysis results (bitmap format)



Clustering using 5 classes

Conclusions

- Remote sensing and GIS are useful tools for the study of spatial and temporal variability of drought indices.

- RDI indentifies the dry and wet events over Greece.

- In the present paper the application of double orthogonal transformation using PCA and Clustering was applied to time series annual RDI index measurements from 1981 to 2001.
- The aim was to define a cost effective, fast and reliable approach using commercial of the shelf software package algorithms that are available to every scientist, stakeholder or public to map and categorize the spatial and diachronic distribution of RDI drought index.
- In conclusion the last approach is considered as the most completed and representative for the Hellenic territory regarding the categorization of drought for the time series data sets ranged from 1981 to 2001 hydrological years.

RELEVANT PUBLICATIONS

1. Kanellou, E., C. Domenikiotis, A. Blanta, E. Hondronikou and N. R. Dalezios, (2008b). Intercomparison of drought indices in semi-arid areas of Greece using conventional data. Proceedings of the International Symposium of Water Shortage Management, Athens, 20 June 2008, pp. 167-179.
2. Kanellou E, Domenikiotis C, Blanta A, Hondronikou E, Dalezios N (2008c) Index-based Drought Assessment in Semi-Arid Areas of Greece based on Conventional Data. European Water (EWRA) 23/24, pp. 87-98.
3. Kanellou E, Tsiros E, Domenikiotis C, Dalezios N (2008d) Drought monitoring using several indices. 4th International Conference on Information and Communication Technologies in Bio and Earth Sciences HAICTA 2008, 18-20 September 2008, Athens, Greece, pp. 32-37.
4. Efrosyni C. Kanellou, Nicos V. Spyropoulos†, Nicolas R. Dalezios, 2011. Geoinformatic Intelligence Methodologies for Drought Spatiotemporal Variability in Greece. Water Resources Management (accepted).

Thank you for your attention

ΕΥΧΑΡΙΣΤΩ

dalezios@uth.gr

Lab. Agromet. University of Thessaly