Research Article

Terrain Segmentation of Greece Using the Spatial and Seasonal Variation of Reference Crop Evapotranspiration

Vassilis Aschonitis, George Miliareis, Kleoniki Demertzis, and Dimitris Papamichail

1Department of Life Sciences and Biotechnology, University of Ferrara, Via L. Borsari 46, 44121 Ferrara, Italy
2Environmental Conservation & Management, Faculty of Pure and Applied Sciences, Open University of Cyprus, Latsia, P.O. Box 12794, 2252 Nicosia, Cyprus
3Department of Hydraulics, Soil Science and Agricultural Engineering, Aristotle University of Thessaloniki, University Campus, 54124 Thessaloniki, Greece

Correspondence should be addressed to Vassilis Aschonitis; schvls@unife.it

Received 13 May 2015; Accepted 8 September 2015

Academic Editor: Jorge E. Gonzalez

Copyright © 2016 Vassilis Aschonitis et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The study presents a combination of techniques for integrated analysis of reference crop evapotranspiration ($\text{ET}_o$) in GIS environment. The analysis is performed for Greece and includes the use of (a) ASCE-standardized Penman-Monteith method for the estimation of 50-year mean monthly $\text{ET}_o$, (b) cross-correlation and principal components analysis for the analysis of the spatiotemporal variability of $\text{ET}_o$, (c) $K$-means clustering for terrain segmentation to regions with similar temporal variability of $\text{ET}_o$, and (d) general linear models for the description of $\text{ET}_o$ based on clusters attributes. Cross-correlation revealed a negative correlation of $\text{ET}_o$ with both elevation and latitude and a week positive correlation with longitude. The correlation between $\text{ET}_o$ and elevation was maximized during the warm season, while the correlation with latitude was maximized during winter. The first two principal components accounted for the 97.9% of total variance of mean monthly $\text{ET}_o$.

$K$-means segmented Greece to 11 regions/clus ters. The categorical factor of cluster number together with the parameters of elevation, latitude, and longitude described satisfactorily the $\text{ET}_o$ through general linear models verifying the robustness of the cluster analysis. This research effort can contribute to hydroclimatic studies and to environmental decision support in relation to water resources management in agriculture.

1. Introduction

The evapotranspiration rate, which can be achieved under no water restrictions from a well-watered ideal grass surface, is called reference crop evapotranspiration $\text{ET}_o$ and it is one of the most important hydroclimatic parameters for the implementation of various hydrological and agricultural applications [1]. Various methods have been developed for the assessment of $\text{ET}_o$ [2, 3] and the ASCE-standardized Penman-Monteith method, which is an update of the FAO-56 Penman-Monteith, has been proposed by the ASCE-EWRI Task Committee as the most precise method for $\text{ET}_o$ estimations [2].

The intra-annual and spatial variability of the $\text{ET}_o$ for regions of Greece has been investigated using different methods by many authors [4–18]. The latest and the most detailed mean monthly $\text{ET}_o$ estimations at country scale based on FAO Penman-Monteith method [1] have been performed by Dalezios et al. [5] using 66 meteorological stations covering a period of 15 years. The authors used the kriging technique for the development of 50 km resolution grids. Mardikis et al. [11] tested different interpolation methods for $\text{ET}_o$ based on 93 meteorological stations and provided improved methods which include the effect of elevation.

The last years, the use of climatic models succeeded to generate high resolution grids of various climatic parameters which enabled the development of global $\text{ET}_o$ maps. Significant works of long-term mean $\text{ET}_o$ estimations at global scales have been performed by many authors [19–22]. $\text{ET}_o$ maps using the FAO-56 Penman-Monteith and ASCE-standardized Penman-Monteith method, which are the most demanding methods in terms of climatic data and
elaboration, have been developed at resolutions of 10 arc-min [20] and 0.5 degrees [21] (1 degree = 60 arc-min = 3600 arc-sec = 1 km at equator). The highest spatial resolution ET$_o$ maps (30 arc-sec) have been developed [22] using the simple method of Hargreaves [23], which requires only temperature data.

Grids of mean monthly ET$_o$ can be used in an attempt to capture the spatial and seasonal patterns and to segment the terrain to regions with distinct seasonal ET$_o$ variability. Common approaches for such analysis are the typical multivariate methods such as correlation analysis, principal components, factorial kriging analysis, multi-Gaussian co-kriging, regression analysis, and so forth, which have successfully been used for other environmental quality parameters [24–30]. In addition, terrain segmentation techniques have been applied to multitemporal land surface temperature (LST) data in an attempt to define subregions with different seasonal LST variability, to assess its sensitivity to climatic change by revealing thermal anomalies and to support environmental analysis [31–36]. Similar regionalization and segmentation approaches have also been applied successfully in order to define subregions with different seasonal precipitation patterns [37–39].

Objectives of the study are (a) to develop high resolution grids (30 arc-sec) of 50-years mean monthly ET$_o$ for Greece using the ASCE-standardized Penman-Monteith, (b) to analyze the spatiotemporal variability of ET$_o$, and (c) to provide a terrain segmentation scheme for Greece based on the spatiotemporal variation of ET$_o$. The techniques of cross-correlation, principal components (PCA), cluster analysis, and general linear models (GLMs) were used. Cross-correlation and PCA are used to capture the dependence of spatiotemporal variation of mean monthly ET$_o$ on geographical attributes (latitude and longitude), topography (elevation), and seasonality. K-means clustering is used to segment the Greek territory to regions/clusters in which the mean monthly ET$_o$ observations approximate to representative mean values with similar temporal variability. GLMs are used to (a) to parameterize the effect of clusters and together with elevation, latitude, and longitude to build models which can provide estimations of mean monthly ET$_o$ (b) to verify the robustness of the cluster analysis and (c) to assess the contribution/effet of each cluster in the mean monthly ET$_o$ estimations. This effort will be valuable in assisting environmental decision support and agricultural planning in relation to water resources management.

2. Data and Methods

2.1. Study Area and Data. The study area is Greece (South-Eat Europe) which is confined between the 34° and 42° parallel N., with a meridional extent from 19° to 28° E. Greece has a typical Mediterranean climate: relatively cold and rainy winters, relatively warm and dry summers, and, generally, extended periods of sunshine. The spatial heterogeneity of climate is mainly attributed to the mountain range of Pindos located in the central part of Greece, while significant differences in the winter’s severity are observed between the central-north continental territory and the islands where

in the second case the winter is milder (Hellenic National Meteorological Service: http://www.hnms.gr/hnms/english/climatology/climatology.html).

The analysis was based on climatic data, which were obtained from the following databases:

(i) [40]: this database provides mean monthly values for the parameters of maximum, minimum, and mean temperature at 30 arc-sec spatial resolution (~1 × 1 km). The data are provided as grids of mean monthly values of the period 1950–2000 (http://www.worldclim.org/). The database also includes a revised version of the GTOP030 DEM based on SRTM DEM at 30 arc-sec spatial resolution. The data were produced using observation-based datasets after elaboration with the thin-plate smoothing spline algorithm implemented in the ANUSPLIN package for interpolation, using latitude, longitude, and elevation as independent variables.

(ii) [41]: this database provides mean monthly values of parameters such as solar radiation, specific humidity, wind speed at 10 m above surface (which was converted to 2 m above surface), precipitation, and temperature for the period 1948–2006 at 0.5 degrees spatial resolution. The dataset was constructed by combining a suite of global observation-based datasets and it is available in the form of NetCDF files (Network Common Data Form) of monthly values of each year for the period 1948–2006 (http://hydrology.princeton.edu/data.pgffield). Resampling to 30 arc-sec spatial resolution was performed using the bilinear resampling scheme. The data were converted to grids of mean monthly values of the period 1950–2000.

2.2. Reference Crop Evapotranspiration. The estimation of ET$_o$ using the ASCE-standardized Penman-Monteith method is performed by the following equation [2]:

\[
ET_o = \frac{0.408 \Delta (R_n - G) + \gamma u_2 (e_s - e_a) C_n / (T_{mean} + 273.16) \Delta + \gamma (1 + C_d u_2)}{1},
\]

where ET$_o$ is the daily reference crop evapotranspiration (mm d$^{-1}$), R$_n$ is the daily net radiation at the crop surface (MJ m$^{-2}$ d$^{-1}$), u$_2$ is the mean daily wind speed at 2 m height above the soil surface (m s$^{-1}$), T$_{mean}$ is the daily mean air temperature (°C), G is the daily soil heat flux density at the soil surface (MJ m$^{-2}$ d$^{-1}$), e$_s$ is the mean daily saturation vapour pressure (kPa), e$_a$ is the mean daily actual vapour pressure (kPa), $\Delta$ is the slope of the saturation vapour pressure-temperature curve (kPa °C$^{-1}$), $\gamma$ is the psychometric constant (kPa°C$^{-1}$), $C_n$ and $C_d$ are constants, which vary according to the time step and the reference crop type and describe the bulk surface resistance and aerodynamic roughness. The short reference crop corresponds to clipped grass of 12 cm height and surface resistance of 70 s m$^{-1}$ where the constants $C_n$ and $C_d$ have the values 900 and 0.34, respectively.
The tall reference crop corresponds to full cover alfalfa of 50 cm height and surface resistance of 45 s m^{-1}, where the constants $C_m$ and $C_d$ have the values 1600 and 0.38, respectively [2]. In this study, the $ET_o$ is estimated using (1) for the commonly used short reference crop as follows: (a) mean monthly values of maximum, minimum, and mean temperature were obtained from the database of [40], while (b) wind speed, specific humidity, and solar radiation were obtained from the database of [41]. The specific humidity was converted to actual vapour pressure [42] before its use. The equations used for intercalculations in ASCE method are given in [2, 43, 44]. All the calculations were performed in ArcGIS 9.3 ESRI environment. Twelve rasters (one for each month) of mean monthly $ET_o$ plus one of mean annual $ET_o$ of the period 1950–2000 were developed for the Greek territory.

2.3. Methodology for Terrain Segmentation Based on the Annual and Seasonal Variation of a Hydroclimatic Parameter

A set of four statistical methods such as cross-correlation analysis, principal component analysis, $K$-means cluster analysis, and general linear models [45–48] are used in this study. Correlation analysis and principal component analysis (PCA) reveal the temporal and spatial pattern evident within a multitemporal dataset. PCA is a linear transformation technique that produces a set of images known as principal components PCs that are uncorrelated with one another while they are ordered in terms of the amount of variance they explain from the original image set [49, 50]. PCs are computed from the linear combination of eigenvectors and the corresponding pixel values of the initial images [51].

PCA has traditionally been used in remote sensing as a means of data compaction since it is usually observed that the first 2 or 3 components are able to explain the majority of the variability in data values. Later components tend to be dominated by noise effects. By rejecting these later components, the volume of data is reduced, with no appreciable loss of information [52]. Standardized principal components analysis [53] is applied (data per month is centered with mean 0.0 and standard deviation 1.0) and so each image is not weighed according to its variance.

$K$-means cluster analysis was used to partition the multitemporal (12-dimensional) imagery of $ET_o$ into $K$ exclusive clusters. It begins by initializing cluster centroids, then assigns each pixel to the cluster whose centroid is nearest, updates the cluster centroids, and then repeats the process until the stopping criteria are satisfied [51]. The analysis uses Euclidian distance for calculating the distances between pixels and cluster centroids. The underlying idea of cluster analysis is that the cluster centroids represent the mean expression of the derived clusters. So clustering of the multitemporal data sets is expected to define groups of pixels with a rather common centroid curve that expresses their average monthly variability [32]. Elevation, latitude, and longitude statistics per cluster centroid curve that express their average monthly variability are computed in order to assist interpretation. Finally the clusters were interpreted according to their centroid and their spatial arrangement [36].

General multiple linear regression analysis was performed using Statgraphics Centurion software (StatPoint Technologies). The parameterization of segmentation effects was made using an integer cluster number $CN$ ranging from 1 to $K$. The $CN$ categorical factor together with the quantitative factors of latitude (Lat), longitude (Lon), and elevation ($H$) were used as independent variables to describe the mean monthly values of $ET_o$. The values of the aforementioned parameters were extracted from 29765 randomly selected positions homogeneously distributed in the entire Greek territory. The selection was performed using the Sampling Design Tool developed by NOAA (National Oceanic and Atmospheric Administration of USA) based on the procedure of “stratified random sampling.” For each month, the mean values of the dependent variable (mean monthly $ET_o$) of sampling positions belonging to each cluster were compared using ANOVA-LSD. This analysis was performed in order to verify that the sampling procedure retained the differences between clusters, which were derived by $K$-means using the whole number of pixels of each cluster. Square root transformation was used for the Lat, Lon, and $H$ values while BoxCox transformation [54] was used for the mean monthly $ET_o$ values to avoid/reduce normality deviations of the dependent and independent variables. The selection of the parameters for the regression model was performed using their variance inflation factor (VIF) in order to avoid multi-co-linearity effects. VIF values above 10 are usually considered to indicate serious multicollinearity [45]. Thus, the following linear model was built to describe the mean monthly $ET_o$ (mm) for each month:

$$
\text{BoxCox}(ET_o, i) = a_0 + \sum_{j=1}^{N} (a_j \cdot I_{\text{CN}, j}) + a_{N+1} \cdot \sqrt{\text{Lon}} + a_{N+2} \cdot \sqrt{\text{Lat}} + a_{N+3} \cdot \sqrt{H},
$$

where $a_0$ is intercept of the regression model and $a_j$ is the regression coefficients of the categorical factor $I_{\text{CN}}$ which is regulated by the cluster number $CN$, $a_{N+1, N+2, N+3}$: the regression coefficients of the quantitative factors Lon, Lat (decimal degrees), and $H$ (m), respectively, and $i$ is the month (starting with $i = 1$ for January and ending with $i = 12$ for December). The value of $N$ is equal to the number of clusters minus 1 ($N = K - 1$) whereas the categorical factor $I_{\text{CN}}$ is regulated by the cluster number $CN$ as follows:

$$
I_{\text{CN}, j} = \begin{cases} 
1, & \text{if } CN = j, \\
-1, & \text{if } CN = N + 1 = K, \\
0, & \text{if } CN \neq j, K.
\end{cases}
$$

The general form of BoxCox transformation for the dependent variable $ET_o$ for each month $i$ is given according to the following formula [54]:

$$
\text{BoxCox}(ET_o, i) = 1 + K_1 \left( (ET_o, i + \lambda_2)^{\lambda_1} - 1 \right)
$$

for $\lambda_1 \neq 0$, where $K_1$ and $\lambda_2$ are constants.
where $\lambda_{1,2}$ is the power and shift parameters of BoxCox transformation, respectively, and $n$ is the number of $ET_o$ observations (samples) for a month $i$. Note that $K_2$ is the geometric mean of $ET_o + \lambda_2$. In order to improve the efficiency of the model, the power parameter $\lambda_1$ was optimized while the shift parameter $\lambda_2$ was set to 0 based on the default optimization procedure incorporated in the statistical software. The optimal transformation is the one that minimizes the mean squared error of the transformed dependent variable [54].

Outliers were not removed during the regression analysis because of the large number of data (there is no automatic procedure to remove outliers). Autocorrelation of the residuals was tested using the Durbin-Watson test [55, 56], which provides values between 0 and 4 where the optimum value for no autocorrelation is equal to 2. ANOVA was used to estimate the statistical significance of the regression model while square correlation coefficient $R^2$ adjusted for degrees of freedom was used to evaluate the explanatory power of the model. Type III sums of squares analysis was used to estimate the statistical significance of the independent variables CN, Lat, Lon, and $H$.

3. Results and Discussion

3.1. Spatial Variation of Mean Annual $ET_o$. The 30 arc-sec resolution (~1 × 1 km) maps of elevation $H$ and mean reference crop evapotranspiration $ET_o$ for the period 1950–2000 estimated by the ASCE standardized Penman-Monteith method are given in Figures 1(a) and 1(b), respectively, while the frequency distributions of pixel values are given, respectively, in Figures 2(a) and 2(b).

The $ET_o$ map (Figure 1(b)) indicates that the lowland areas (below 250 m) with a distance less than 20–25 km from the shoreline and with latitude less than 39 degrees present annual values of $ET_o > 1200$ mm. Regions with extremely high mean annual $ET_o$ values >1400 mm occupy approximately 13.5% of the Greek territory and they are distributed mainly to lowlands along the coastlines of (a) Crete and Dodecanese islands (southern and eastern Aegean Sea), (b) eastern Attika, and (c) eastern Peloponnesus. These findings are in accordance to the study of [5] who used a 50 km resolution grid and the studies of [57, 58] who analyzed data from meteorological stations.

3.2. Spatial and Seasonal Variation of Mean Monthly $ET_o$. The correlation matrix of the mean monthly values of reference evapotranspiration $ET_o$ versus $H$, Lat, and Lon is quantified in Table 1. According to Table 1, the mean monthly $ET_o$ presents negative correlations with $H$ and Lat while a rather weak positive correlation between $ET_o$ and Lon exists. The absolute values of $ET_o$-$H$ correlations are maximized during the warm season while the $ET_o$-Lat correlations are

![Figure 1: Maps of (a) elevation $H$ and (b) mean annual $ET_o$ for the period 1950–2000 estimated by the ASCE standardized Penman-Monteith method.](image)
maximized during the winter. The absolute values of ET_o-Lon correlations are generally low and they are maximized during the period of August–January.

PCA was implemented in order to interpret the monthly intercorrelations observed in Table 1 and the corresponding eigenvalues and eigenvectors are presented in Table 2. According to the eigenvalues, the first two principal components (PC-1 and PC-2) account for the 97.9% of the total variance observed in the 12 monthly ET_o images.

The PC-1 map (Figure 3(a)) is composed of linear combinations of monthly ET_o images which are not influenced by seasonality since the weights of eigenvectors slightly vary between 0.28 and 0.30 (Table 2). Taking into account the correlation coefficients of monthly ET_o versus H and Lat from Table 1, it is easy to detect that PC-1 amplifies the effects of elevation and latitude. The effects of elevation and latitude indirectly express the effects of temperature, which is higher in the lowlands but also higher in lower latitudes. The PC-1 is spatially maximised in the lowlands of the southern continental Greece and in the islands.

For PC-2 map (Figure 3(b)), the linear combinations of monthly ET_o images amplify the difference between

---

**Figure 2:** Frequency analysis of the spatial variation of (a) elevation $H$ and (b) mean annual ET_o estimated by the ASCE standardized Penman-Monteith method for the period 1950–2000.

**Figure 3:** The first two principal components PC-1 and PC-2 that account for the 97.9% of the ET_o variance.
Table 1: Cross-correlation matrix of the mean monthly ET<sub>ᵦ</sub> (January to December) as well as the latitude, longitude, and elevation images. Correlations are performed between pairs of variables taking values from pixels of the same geographical position.

<table>
<thead>
<tr>
<th>H</th>
<th>Lat</th>
<th>Lon</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>1</td>
<td>0.03</td>
<td>0.14</td>
<td>0.32</td>
<td>0.64</td>
<td>0.57</td>
<td>0.75</td>
<td>0.90</td>
<td>0.98</td>
<td>0.10</td>
<td>0.50</td>
<td>0.42</td>
<td>0.32</td>
<td>0.14</td>
</tr>
<tr>
<td>Lat</td>
<td>0.03</td>
<td>1</td>
<td>0.21</td>
<td>0.97</td>
<td>0.21</td>
<td>0.97</td>
<td>0.21</td>
<td>0.97</td>
<td>0.21</td>
<td>0.97</td>
<td>0.21</td>
<td>0.97</td>
<td>0.21</td>
<td>0.97</td>
</tr>
<tr>
<td>Lon</td>
<td>-0.30</td>
<td>-0.14</td>
<td>1</td>
<td>0.32</td>
<td>0.64</td>
<td>0.57</td>
<td>0.75</td>
<td>0.90</td>
<td>0.98</td>
<td>0.10</td>
<td>0.50</td>
<td>0.42</td>
<td>0.32</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Table 2: Eigenvalues and eigenvectors for the PCs of the mean monthly ET<sub>ᵦ</sub>.

<table>
<thead>
<tr>
<th>Eigenvectors</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal components (PCs)</td>
<td>0.28</td>
<td>-0.41</td>
<td>0.33</td>
<td>-0.19</td>
<td>-0.40</td>
<td>-0.03</td>
<td>0.00</td>
<td>0.04</td>
<td>-0.44</td>
<td>0.33</td>
<td>-0.34</td>
<td>-0.17</td>
</tr>
<tr>
<td>Jan</td>
<td>0.29</td>
<td>-0.32</td>
<td>-0.16</td>
<td>-0.36</td>
<td>0.06</td>
<td>0.06</td>
<td>0.16</td>
<td>-0.52</td>
<td>0.10</td>
<td>0.21</td>
<td>0.50</td>
<td>0.20</td>
</tr>
<tr>
<td>Feb</td>
<td>0.29</td>
<td>-0.20</td>
<td>-0.49</td>
<td>-0.19</td>
<td>0.11</td>
<td>0.61</td>
<td>-0.02</td>
<td>0.22</td>
<td>0.03</td>
<td>-0.28</td>
<td>-0.28</td>
<td>-0.14</td>
</tr>
<tr>
<td>Mar</td>
<td>0.29</td>
<td>0.05</td>
<td>-0.46</td>
<td>0.39</td>
<td>-0.14</td>
<td>-0.39</td>
<td>0.09</td>
<td>-0.43</td>
<td>-0.09</td>
<td>-0.12</td>
<td>-0.39</td>
<td>0.11</td>
</tr>
<tr>
<td>Apr</td>
<td>0.29</td>
<td>-0.25</td>
<td>-0.25</td>
<td>0.19</td>
<td>-0.58</td>
<td>0.04</td>
<td>-0.26</td>
<td>0.26</td>
<td>-0.06</td>
<td>0.06</td>
<td>0.51</td>
<td>-0.16</td>
</tr>
<tr>
<td>May</td>
<td>0.29</td>
<td>0.31</td>
<td>-0.08</td>
<td>-0.30</td>
<td>0.07</td>
<td>-0.31</td>
<td>0.66</td>
<td>0.39</td>
<td>0.11</td>
<td>0.13</td>
<td>-0.03</td>
<td>-0.09</td>
</tr>
<tr>
<td>Jun</td>
<td>0.28</td>
<td>0.44</td>
<td>0.11</td>
<td>-0.46</td>
<td>0.00</td>
<td>-0.05</td>
<td>-0.47</td>
<td>-0.02</td>
<td>0.01</td>
<td>-0.04</td>
<td>-0.21</td>
<td>0.48</td>
</tr>
<tr>
<td>Jul</td>
<td>0.28</td>
<td>0.35</td>
<td>0.34</td>
<td>-0.04</td>
<td>0.11</td>
<td>0.12</td>
<td>-0.01</td>
<td>-0.44</td>
<td>0.02</td>
<td>-0.25</td>
<td>0.01</td>
<td>-0.63</td>
</tr>
<tr>
<td>Aug</td>
<td>0.29</td>
<td>0.17</td>
<td>0.34</td>
<td>0.46</td>
<td>0.02</td>
<td>0.49</td>
<td>0.33</td>
<td>-0.01</td>
<td>-0.06</td>
<td>0.10</td>
<td>0.00</td>
<td>0.44</td>
</tr>
<tr>
<td>Sep</td>
<td>0.30</td>
<td>-0.04</td>
<td>-0.08</td>
<td>0.21</td>
<td>0.67</td>
<td>-0.16</td>
<td>-0.28</td>
<td>0.18</td>
<td>-0.39</td>
<td>0.28</td>
<td>0.16</td>
<td>-0.10</td>
</tr>
<tr>
<td>Oct</td>
<td>0.29</td>
<td>-0.24</td>
<td>0.15</td>
<td>0.20</td>
<td>-0.02</td>
<td>-0.10</td>
<td>-0.23</td>
<td>0.11</td>
<td>0.78</td>
<td>0.27</td>
<td>-0.16</td>
<td>-0.10</td>
</tr>
<tr>
<td>Nov</td>
<td>0.28</td>
<td>-0.36</td>
<td>0.28</td>
<td>0.04</td>
<td>0.01</td>
<td>-0.29</td>
<td>0.01</td>
<td>0.21</td>
<td>-0.03</td>
<td>-0.71</td>
<td>0.21</td>
<td>0.16</td>
</tr>
<tr>
<td>Dec</td>
<td>11.13</td>
<td>0.62</td>
<td>0.15</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Variance %</td>
<td>92.71</td>
<td>5.17</td>
<td>1.25</td>
<td>0.32</td>
<td>0.18</td>
<td>0.12</td>
<td>0.08</td>
<td>0.06</td>
<td>0.04</td>
<td>0.04</td>
<td>0.02</td>
<td>0.02</td>
</tr>
</tbody>
</table>

the periods of October–March and April–September which present negative and positive weights of eigenvectors, respectively (Table 2). The difference between these two periods is spatially maximised in central Greece, in central Macedonia (Thessaloniki plain and Chalkidiki regions), in eastern Thrace, in the islands of north-eastern Aegean sea and south Ionian sea. These patterns are in accordance with the spatiotemporal patterns given by [5, 8, 59].

3.3. Terrain Segmentation Based on the Spatial and Seasonal Variation of ET<sub>ᵦ</sub>. For the implementation of the K-means method, the following criteria were used: (a) small clusters with area extent (occurrence) less than 0.001% were eliminated by merging them with larger clusters that are closest to their centroids and the stopping criterion was defined as the percentage of the migrating pixels during a specific iteration (if it was less than 0.001% of the entire image pixels the clustering was terminated), (b) the maximum number of clusters K was defined by a trial and error procedure, and (c) the maximum number of iterations was set equal to 300. At the end of the analysis, the mean monthly values of ET<sub>ᵦ</sub> are given for each cluster (described as multipart polygons).

Cluster analysis revealed 11 clusters (K = 11) in which the mean monthly ET<sub>ᵦ</sub> observations approximate to representative mean values with similar temporal variability. The characteristics of their centroids (latitude, longitude), their percent area extent, mean elevation, and the variation of monthly ET<sub>ᵦ</sub> are given in Table 3. The spatial distribution of the 11 clusters is given in Figure 4.
In general, \(\text{ET}_o\) is maximized during summer and minimized during winter where the 76.7% of annual \(\text{ET}_o\) in Greece is observed during April to September (warm-dry period), while the 23.3% is observed during October to March (cold-wet period) (Table 3). Clusters of significant interest are the following:

(i) Clusters 9, 10, and 11 present mean annual values of \(\text{ET}_o > 1400\) mm, cluster 9 presents the highest mean monthly \(\text{ET}_o\) among all clusters during the period of May–September, and cluster 11 presents the highest mean monthly \(\text{ET}_o\) among all clusters during the period of October–April.

(ii) Clusters 3 and 4 present the lower mean annual values of \(\text{ET}_o\) and correspond to the higher elevation areas of Greece.

(iii) Clusters 1 and 5 cover the major part of the agricultural land in Greece.

### 3.4. Regression Models of Mean Monthly \(\text{ET}_o\)

The eleven clusters (\(K = 11\)) obtained by the \(K\)-means analysis were used to calibrate (2) for each month using the observations of monthly \(\text{ET}_o\), Lon, Lat, \(H\), and CN values of the 29765 selected sampling positions. The sampling positions were distributed to each cluster based on the % cluster’s area extent (Table 3). Their number within each cluster is given in Table 4. The comparison of the mean monthly values of the dependent variable (the BoxCox transformed mean monthly \(\text{ET}_o\)) of sampling positions belonging to each cluster based on ANOVA-LSD is also given in Table 4. This analysis verified that the sampling procedure retained the differences between clusters which were derived by \(K\)-means taking into account the fact that only 4 out of 660 combinations of pairs comparisons were found nonstatistically different (e.g., pairs of clusters 3 and 4 for January, clusters 9 and 10 in June, clusters 7 and 9 in August, and clusters 9 and 10 in October) (Table 4).

The variance inflation factors VIFs of the selected independent variables of (2) ranged between 1.26 and 3.3 (the maximum VIF value corresponds to the transformed value of elevation). The regression coefficients and the statistical tests of general multiple regression analysis are given in Table 5 (twelve models, one for each month) whereas the respective graphs of observed versus predicted values of the BoxCox
transformed mean monthly $\text{ET}_o$ are given in Figure 5. For all models, type III sums of squares analysis showed that the independent variables CN, Lat, Lon, and $H$ were statistically significant at 95% confidence level with $p < 0.001$, while the Durbin-Watson test indicated no serial autocorrelation of the residuals at 95% confidence level. Square correlation coefficients $R^2$ ranged between 0.88 and 0.96 indicating satisfactory accuracy of the twelve models.

The analysis of the general linear models verified the robustness of the cluster analysis in two ways (a) by the statistically significant difference between the clusters, which was derived by the ANOVA-LSD (Table 4), and (b) by the high statistical significance of the CN parameter as independent variable in the models and their high square correlation coefficients. The contribution/effect of each cluster in the spatiotemporal variation of $\text{ET}_o$ in the Greek territory, as it
Figure 5: Continued.
Figure 5: Observed versus predicted values of the BoxCox transformed mean monthly ET, using the models of Table 5.
is described by the sum term inside (2), is given in Figure 6. According to Figure 6, clusters 1, 2, 3, and 4 present a negative effect while clusters 6, 7, 9, 10, and 11 present a positive effect to all months. The effect of clusters 5 and 8 presents small positive values during the warm season and small negative values during the cold season. The contribution of each cluster (Figure 6) defines the relative behaviour of each cluster during each month using, as a reference, the mean monthly conditions of ET₀ in the Greek territory.

4. Conclusions
The derived high resolution dataset of mean monthly ET₀ and the statistical analysis captured the seasonal and spatial variation of ET₀ in Greece while the regional spatiotemporal trends were revealed and assigned to spatial clusters. More specifically, cross-correlation and principal component analysis quantified the temporal pattern of ET₀ and revealed season, latitude, longitude, and elevation dependencies, while
$K$-means cluster analysis segmented the terrain to regions with distinct temporal variability of $\text{ET}_o$. The derived spatiotemporal patterns are in accordance with previous research efforts.

The parameterization of each cluster (revealing a unique temporal signature of monthly $\text{ET}_o$) together with topography (elevation) and spatial arrangement (latitude and longitude) described satisfactorily the mean monthly $\text{ET}_o$ through generalized linear models verifying the robustness of the cluster analysis, providing also a tool for $\text{ET}_o$ estimations for agricultural purposes. The feasible development of high resolution grids of monthly $\text{ET}_o$ based on available meteorological data produced by climatic models and the terrain segmentation approach can be used as tools to design decision support and management schemes of water recourses.

The procedures described in this study can be used as a general framework for the spatial and seasonal variation not only for $\text{ET}_o$ but also for other climatic parameters while it can support future research efforts such as

(a) assessing any climate changes by integrating monthly $\text{ET}_o$ data sets or other climatic variables of the forthcoming years; the new data sets can be recomputed the temporal pattern of the climatic variable within each cluster in an attempt to evaluate any trends in the centroid change between two periods;

(b) using the proposed framework at watershed scale in order to facilitate the surface integration of climatic variables obtained by meteorological stations based on clusters produced by long-term data such as the ones of this study.

**Conflict of Interests**

The authors declare that there is no conflict of interests regarding the publication of this paper.

**References**


[15] F. Xystrakis and A. Matzarakis, “Evaluation of 13 empirical reference potential evapotranspiration equations on the island...


Submit your manuscripts at http://www.hindawi.com