HYDROLOGICAL PROCESSES
Hydrol. Process. (2012)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/hyp.9507

Pixel-oriented land use classification in energy balance modelling

Giuseppina Scavone,[1]* Juan Manuel Sánchez,[2†] Vito Telesca,[1] Vicente Caselles,[2] Vito A. Copertino,[1] Vittoria Pastore[1] and Enric Valor[2]

[1] *Department of Environmental Engineering and Physics, University of Basilicata, Viale dell'Ateneo Lucano, 10-85100 Potenza, Italy*
[2] *Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia, Dr. Moliner, 50, 46100 Burjassot, Spain*

Mass and energy transfer between soil, vegetation and atmosphere is the process that allows to maintain an adequate energy and water balance in the earth–atmosphere system. However, the evaluation of the energy balance components, such as the net radiation and the sensible and latent heat fluxes, is characterized by significant uncertainties related to both the dynamic nature of heat transfer processes and surfaces heterogeneity. Therefore, a detailed land use classification and an accurate evaluation of vegetation spatial distribution are required for an accurate estimation of these variables. For this purpose, in the present article, a pixel-oriented supervised classification was applied to obtain land use maps of the Basilicata region in Southern Italy by processing three Landsat TM and ETM+ satellite images. An accuracy analysis based on the overall accuracy index and the agreement Khat of Cohen coefficient showed a good performance of the applied classification methodology and a good quality of the obtained maps. Subsequently, these maps were used in the application of a simplified two-source energy balance model for estimating the actual evapotranspiration at a regional scale. The comparison between the simulations made by applying the simplified two-source energy balance model and the measurements of evapotranspiration at a lysimetric station located in the study area showed the applicability and the validity of the proposed methodology. Copyright © 2012 John Wiley & Sons, Ltd.

KEY WORDS pixel-oriented land use classification; energy balance; actual evapotranspiration; heat fluxes; Landsat

Received 30 July 2011; Accepted 4 July 2012

INTRODUCTION

In the last years, Mediterranean regions have been affected by a significant reduction in rainfall and have consequently become aware on drought and desertification problems, especially in regions characterized by arid or semiarid climate. The issue of water scarcity is becoming more pressing also for the geographic areas that had never experienced the phenomenon until now. This is a direct consequence of the considerable changes in rainfall against the growing increase in the exploitation of water resources due to the rising civil, agricultural and industrial demand. According to the international organizations, the risk of desertification is becoming more and more tangible in Mediterranean countries, and the actions to be taken to mitigate this soil degradation process can be defined according to several lines, depending on the different available scientific expertise, and in a global approach, overcoming the sectoral analyses and covering a properly calibrated territory. Therefore, a continuous monitoring of the phenomenon, of its hydro-meteorological and climatic forcing (such as the evapotranspiration, object of the present work) and of its effects, also using the methods and the innovative technologies offered by the recent progresses in scientific research, is essential for preventive and control purpose.

In this context, the actual evapotranspiration is one of the main variables to estimate because it is a key element of many physical processes that occur in the atmosphere. A proper formulation and parameterization of this phenomenon is crucial for the development and improvement of hydrological, meteorological and climatologic models. However, it is also very difficult to estimate, especially in areas characterized by the presence of sparse vegetation. Several studies showed that the level of detail in vegetation cover characterization represents a fundamental element in the evaluation of soil–vegetation–atmosphere heat fluxes, and it may have a remarkable influence on the accuracy of evapotranspiration estimate (Moran *et al.*, 1997; Kustas and Norman, 1999; Humes *et al.*, 2002; Valor and Caselles, 2005; Sánchez *et al.*, 2007, 2008a).

The survey activity carried out using sensors placed on satellite platforms represents a useful tool in this sense and a valid support, complementary and not substitutive, to traditional field surveys. In fact, the remote sensing techniques allow to obtain, at a global covering, a reliable evaluation of critical parameters for the evapotranspiration retrieval, such as the surface thermodynamic temperature, that can be obtained from the brightness temperature direct sensed by the remote instruments, the net radiation, from a balance equation between long-wave and short-wave radiations, the sensible heat flux as a combination of vegetation and soil contributions. Instead, field data are used to estimate the typical parameters of the atmosphere evaporating power, such as wind speed, global radiation and air temperature. The integration of remote sensing data with field data allows the application

*Correspondence to: Giuseppina Scavone, Department of Environmental Engineering and Physics, University of Basilicata, Viale dell'Ateneo Lucano, 10-85100 Potenza, Italy.
E-mail: scavone@unibas.it
†Present address: Department of Applied Physics, University of Castilla-La Mancha, Plaza de Manuel Meca, 1-13400 Almadén, Spain.

of actual evapotranspiration models based on the energy balance equation.

The objective of the present work was to verify the applicability of a methodology for the daily actual evapotranspiration retrieval, at a regional scale, adopting an approach based on a pixel-oriented supervised land use classification obtained by Landsat TM and ETM+ images within a simplified two-source energy balance (STSEB) model. The STSEB model was extensively analyzed, applied and validated in literature by several authors (Copertino *et al*., 2006; Sánchez *et al*., 2007, 2008a, b, c) over a variety of surface conditions and at different spatial scales, with good results. The present work focuses on the possibility of improving the STSEB model performance through a very accurate estimation of aerodynamic resistances derived from a regional detailed land use spatial reconstruction developed through a 'pixel-oriented supervised classification' by processing the same Landsat TM and ETM+ satellite images later used to estimate the actual evapotranspiration.

In particular, the land use classification procedure and the STSEB model were applied to the Basilicata region (South Italy), through the integration of field data, derived from agro-meteorological stations distributed in the study area and Landsat remote sensing data.

The work was organized as follows. The first phase consisted in processing Landsat images to realize land use thematic maps of the study area (Basilicata region, Southern Italy) at a regional scale, by selecting a pixel-oriented supervised classification methodology. The STSEB model was then applied for actual evapotranspiration estimate using the regional thematic maps of land use to characterize the distribution of vegetation and to evaluate the contributions of soil and vegetation to the total sensible heat flux. An equilibrium relationship between short-wave and long-wave radiations was applied to estimate the instantaneous net radiation. The actual evapotranspiration was finally obtained as a residual term of the energy balance equation, once the net radiation and the sensible heat flux were estimated. The results were converted from instantaneous to daily scale through the R_{nd}/R_{ni} coefficient (ratio between daily, R_{nd}, and instantaneous, R_{ni}, net radiations) and compared with the evapotranspiration measured by a weighing lysimeter located in the Ionian band of Basilicata region.

STUDY AREA

The Basilicata is the most mountainous region of Southern Italy, with 47% of its area (~10000 km^2) covered by mountains, 45% made up of hills and only 8% made up of plains. Two main zones can be distinguished: the western one, formed by the Appennino Lucano, about coincident with the province of Potenza, and the eastern hilly area that forms essentially the province of Matera and that declines to the coastal plain on the Ionian Sea, along which stretches the only real plain of the region. The short coastline (~15 km) along the Tyrrhenian Sea is high and rocky, whereas approximately 35 km of shore, on the Ionian coast, from Metaponto to Nova Siri, are characterized by predominantly sandy beaches, where five of the main Basilicata rivers flow to the sea: Bradano, Basento, Cavone, Agri and Sinni. The cultivated areas of the Metaponto plain, which extend behind the coastal band, bordering to the southeast the Ionian Sea and behind the declining hills of Matera, rich in historical and natural resources, represent a real treasure of the regional agriculture, whose productions of fruit and vegetables are exported all over Europe.

The Ionian band is characterized by a warm temperate Mediterranean climate. The average annual temperature varies from 14.5 °C to 16.9 °C, whereas the monthly average temperature is colder, varying between 6 °C and 9.9 °C. Rainfall is very irregular but, overall, not scarce; generally, it increases proceeding from South to North and from East to West, where the high massifs exert a more significant action of winds capture.

The different climatic conditions also correspond to different types and distributions of vegetation cover: dense vegetation, mainly consisting in deciduous and coniferous woods, characterizes the mountain areas to the West, whereas a variety of crops characterizes the central areas as well as the hilly southeastern zone and the plain areas, up to the bare soil that mainly covers the boundary zone with the Puglia region. Therefore, a detailed characterization of land use and land cover appears of great importance for the estimation of physical variables, such as the actual evapotranspiration, at a regional scale.

In the present work, the Ionian zone of the Basilicata region (Figure 1) was chosen as study area due to both socioeconomic and touristic importance that this area has for the region and for the South of Italy and the availability of useful field and remote sensing data for the application and validation of the proposed model.

DATA

The application of STSEB model, in the proposed approach, required the integration of field and remote sensing data.

Field meteorological data were provided by the Agenzia Lucana di Sviluppo e di Innovazione in Agricoltura (ALSIA) stations. They are equipped with instruments for continuous and automatic measurement of the following meteorological parameters:

- temperature (°C), pressure (bar) and relative humidity (%) of the air
- soil temperature (°C)
- precipitations (mm)
- evaporation (mm day^{-1});
- wind direction and speed (km h^{-1})
- solar global radiation (MJ m^{-2})
- leaf wetness (min)

ALSIA agency, activated in 1996, operates a network of 38 agro-meteorological stations distributed throughout the region (Figure 1). The service is organized on two levels:

DOI: 10.1002/hyp

Figure 1. Study area: Ionian band of the Basilicata region (Southern Italy); the red dots indicate ALSIA meteorological stations distributed all over the region

Level I: activities related to the maintenance of the stations network.

Level II: activities related to the quality control of data and their reconstruction, editing of the weekly agro-meteorological bulletin and regional and climatologic analysis and studies.

Table I shows the names of ALSIA stations (ID) located in the study area, the communes to which they belong, their elevation above sea level and geographical coordinates.

In particular, the field data used in this work were global solar radiation, wind speed and air temperature.

At ALSIA stations, air temperature is measured by a PT100 thermoresistor, with an accuracy of $\pm 0.1\,^{\circ}C$; global solar radiation is measured by a radiometer CM7-Kippen & Zonen, composed by two pyranometers, working in the range (0.3–2.5 μm), with an accuracy of ±2%. The instruments for wind speed measurement are classical three-cup anemometers, located at 2 m above the ground, having an accuracy of $\pm 0.5\,ms^{-1}$.

Regional maps of air temperature, wind speed and global solar radiation were created by interpolating the data registered at ALSIA stations to integrate them with the remote sensing data.

Two Landsat_7-ETM+ images, corresponding to 26 September 1999 and 14 June 2002, and a Landsat_5-TM image, corresponding to 26 May 2004, were selected in clear sky conditions. Landsat images were chosen because of their spatial resolution (30 × 30 m), suitable for detecting the variability of vegetation cover, which must be taken into account in applying STSEB model. They were georeferenced, corrected from atmospheric effects and processed both to obtain maps of land use of the Basilicata region by applying a supervised classification, as described in the Pixel-Oriented Land Use Classification section, and to obtain instantaneous net radiation, R_n, radiometric surface temperature, T_s, and sensible heat flux, H, required by STSEB model, as described in the Application of STSEB Model section.

Table I. Meteorological ALSIA stations

ID	Commune	High (m.a.s.l.)	Latitude (N)	Longitude (E)
CRP	Pisticci	51	40.434	16.761
ME3	Bernalda	68	40.414	16.580
MO4	Montalbano	154	40.391	16.790
MO6	Tursi	92	40.364	16.620
NS3	Nova Siri	152	40.363	16.536
PAN	Metaponto	24	40.281	16.615
PI1	Pisticci	192	40.252	16.567
PI3	Pisticci	52	40.226	16.626
PO1	Policoro	117	40.183	16.688
PO3	Policoro	6	40.148	16.589

In particular, for the atmospheric correction of the satellite data, atmospheric profiles from Brindisi radiosound were used in the radiative transfer model MODTRAN 4.0. The radiosounding station of Brindisi belonging to Cloudless Land Atmospheric Radiosounding (CLAR) database is located next to the test area and has homogeneous characteristics to it. The CLAR database is created by atmospheric radiosoundings taken by the Atmospheric Science Department, University of Wyoming (http://weather.uwyo.edu/upperair/sounding.html). It collects data from 382 global land atmospheric radiosoundings uniformly distributed. All data are checked to ensure the absence of clouds. A radiosounding is considered 'cloudy' when the relative humidity (RH) of a layer is higher than 90% or when two successive layers have RH > 85%. A radiosounding with RH higher than 80% in the first two layers is considered 'foggy' and then discarded.

Furthermore, in a cultivated site near the agro-meteorological station of Policoro (40°10′15″N, 16°38′53″E), in the province of Matera, there is a weighing lysimeter that records values of actual evapotranspiration. It consists of a reservoir having an area of 2 × 2 m and 1.30 m deep, isolated from any external structure and filled with a soil having the same characteristics and cultivated with the same crops of the surrounding area. Water requirement and daily evapotranspiration are directly obtained from the difference in weight between two subsequent readings. The measurement accuracy of this system is approximately ±0.06 mm. The 'Istituto di Scienze della Produzione Alimentare di Bari,' in charge of collecting the data coming

DOI: 10.1002/hyp

from the experimental farm 'E. Pantanelli' of Policoro (MT), provided actual evapotranspiration values corresponding to 26 September 1999 on a crop of radish, 14 June 2002 on a crop of tomatoes and 26 May 2004 on a crop of sweet corn. These values were used in the present work as ground data for a detailed comparison with the values extracted from the actual evapotranspiration maps obtained by the application of STSEB model.

PIXEL-ORIENTED LAND USE CLASSIFICATION

The novelty of the present work compared with the previous applications of the STSEB model by the authors consists in having made specific land use maps of the study area, the Basilicata region (South Italy), later used, instead of CORINE land Cover map, for characterizing the vegetation cover and land use distribution. In fact, a detailed land use spatial reconstruction can improve the STSEB model performance through a more accurate estimation of aerodynamic resistances.

The production of thematic maps requires the application of methods for the classification of the territory in areas corresponding to the typologies selected in the legend of the map. During the elaboration phase, different statistic and/or deterministic, analogical and/or digital processes can be identified. In using the information contained in each frequency band to treat rapidly the large amount of data produced by the different survey systems, procedures for the recognition of spectral configurations with special algorithms are preferably applied. First, a choice must be made between the two traditional classification techniques, supervised and unsupervised. The difference is that a preliminary phase of *in situ* human control of data is required in the supervised classification, whereas it is not required in the unsupervised classification. In the case of supervised classification, the acquisition of field reference data is useful both for guiding classifiers and for providing validation information of the final products; in addition, the acquisition of georeferenced data allows to link field and remote sensing data.

In an unsupervised classification, the objective is to group into 'clusters' the multiband spectral responses that are statistically distinguishable. Generally, in an area within the image, a certain number of pixels belonging to the same cluster corresponds to a class or structure to the ground (initially unknown), so that in a new image that displays the spatial distribution of the clusters, the structures will appear highlighted by the distribution of grey levels. Different colours can be assigned to these levels to produce a 'map' of the clusters. The aim is to associate categories of ground objects to different clusters, on the basis of prior knowledge of the principal classes attended in the scene or, if possible, acquiring so-called 'ground truth' by visiting the area, and visually correlating the structures in the map to the corresponding ones on the ground, as there is no prior selection of classes. The input data for an unsupervised classification are the digital number values of different bands, from which the algorithms calculate the clusters values. The maximum number of clusters is automatically determined by the parameters selected in the elaboration phase. The typical result of this operation is to generate many clusters, as numerous as the resulting classified image that appears too complex and difficult to read in terms of assigned classes. The unsupervised technique is particularly useful when analyzing not well-known areas or inaccessible and unexplored ones. This methodology may also be appropriate in exploratory image analysis because it allows to appraise how many classes can be highlighted and how they can be spectrally distinguished.

The supervised classification is particularly efficient in terms of accuracy, when classes whose validity depends largely on human knowledge and experience have to be represented. It is based on knowledge of some ground sample areas representative of the surface classes that will make up the thematic map, well known and conveniently located on the images. The procedure can be summarized as follows:

- thematic classes are defined and the corresponding legend is created;
- sample areas for each class are localized on the classifying images;
- the available bands (belonging to the same survey) less correlated among them, with respect to the selected classes, and those where the spectral differentiation among different explored classes is maximum are chosen;
- the statistical parameters of the selected thematic classes are determined (mean, covariance, maximum and minimum) based on values of pixels belonging to the sample areas to obtain the typical values of radiance for each class, for each selected band;
- finally, the most suitable classification algorithm for assigning each pixel to a specific class is chosen; the most widely used is the 'maximum likelihood'.

The pixel-oriented supervised classification methodology is as follows: conventional (real or well known) or meaningful classes are identified in the scene through *a priori* knowledge due to familiarity with the area of interest. Alternatively, the recognition may take place through thematic maps or field surveys. This allows to define discrete classes to which assigning the names of the identified categories (the so-called 'guided selection'). The sample areas, represented by each known area that appears enough homogeneous on the image (in tone or colour with the typical form of the category), are identified and circumscribed in a polygon on the image displayed on the monitor. For each outlined class, the average values and variances are calculated in digital numbers for each band used in the classification and for all the pixels in the area (more than one polygon is defined for each class). The digital values are then plotted as function of the sequence of bands in order of wavelength. The resulting curve is the spectral signature or spectral response of that class. A statistical processing allows to assign each pixel to the class having the most similar spectral signature. Few pixels cannot be assigned and remain unclassified; they may

DOI: 10.1002/hyp

belong to undefined or unrecognized classes of the available spectral library.

A disadvantage of this approach is that the spectral signatures for different vegetation types in the library can also capture differences in the same class, distinguishing the different elements belonging to it. This distinction involves considerable difficulties for the results analysis since it is complex to distinguish, investigating a satellite image, for instance, in what is recognized as 'woods of broad-leaf or conifers' the particular species that comprises. Much still needs to be done about the best use of this type of study on satellite image.

Land use classification of the Basilicata region

In the present work, to generate land use thematic maps of the Basilicata region, both the supervised and the unsupervised classification techniques were initially applied to the three Landsat_5-TM and Landsat_7-ETM + images corresponding to 26 September 1999, 14 June 2002 and 26 May 2004.

The study of the area was carried on using the following data:

- orthophotos (1:5000) of Balvano, Basento mouth, Marsico Nuovo, North Potenza, Tito Scalo and Cavone mouth for the years 1999 and 2004;
- IGM cartography of the Basilicata (1:25000);
- three Landsat images: one Landsat_5-TM for 26 May 2004 and two Landsat_7-ETM+ for 26 September 1999 and 14 June 2002;
- *in situ* measurements campaigns.

The classification started from a phase of images preprocessing, including procedures for acquisition, georeferencing and geocoding, correction of systematic biases, cleaning, stress and visualization, aimed at assessing and eliminating 'spurious' signals in the data, due to atmospheric effects, spectral noises and sampling errors, to enable the implementation of the classification algorithms. Areas of interest were selected through the available thematic maps and orthophotos. Investigations at different levels were carried on to provide a detailed picture of the vegetation cover and to support further analyses. In that regard, field survey campaigns were conducted (to integrate the information coming from the analysis of orthophotos) to acquire knowledge about vegetation in the study areas.

In the following classification phase, the different soil types and their radiometric and spectral characterization were identified, testing both the supervised and the unsupervised procedure, coming to choose the first. To appraise its reliability, it was applied to some training areas through the comparison with *in situ* observations. Finally, the obtained land use maps were compared with the CORINE Land Cover 2000.

After applying the supervised classification to a selected portion of the images with well-known vegetation and land use characteristics, the study was extended to the entire Basilicata region, through the identification of regions of interest, repeating the procedure of classification (Figure 2).

Among the available classification algorithms, the maximum likelihood, belonging to the supervised category, gave the best results. This classifier assigns each pixel to the class for which the conditional probability is greater. The latter is defined as the probability that, selecting a pixel X of the scene, it belongs to one specific class C_c. The pixel X belongs to the class C_c if

$$P(C_c \backslash X) > P(C_k \backslash X) \quad \text{for each} \quad k <> c \qquad (1)$$

where $P(C_c|X)$ indicates the conditional probability of pixel X in relation to the class C_c, and $P(C_k|X)$ indicates the conditional probability of pixel X in relation to the class C_k.

The conditional probabilities $P(C_c|X)$ are never *a priori*. It is assumed, however, to have an available number of samples for each class to be able to estimate with good accuracy the probability that given a class C_c a pixel of value X is extracted; this probability is designated by $P(X|C_c)$. There will be so many probabilities as many the classes of coverage selected.

In this classification procedure, it is assumed that the distribution of reflection values for each class can be approximated to a multinormal distribution in the multidimensional space of the electromagnetic spectrum.

During the classification, reflectance values $X = (x_1, x_2, \ldots . x_n)$ are read for each pixel of the image; the mean values and the matrixes of variance–covariance (calculated from samples of the m classes) are then considered, which allow to calculate the probabilities $P(X|C_i)$ for each class. The probability $P(X|C_i)$, represented by the multinormal distribution, cuts down only to infinity, which means that

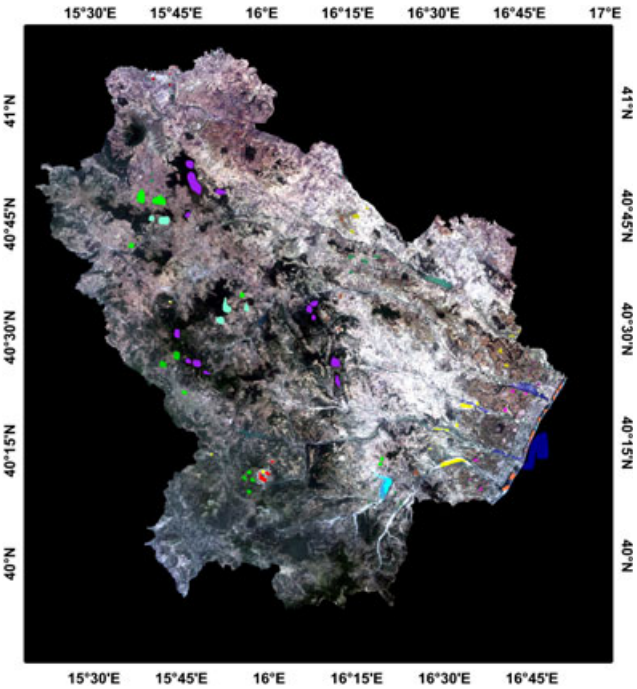

Figure 2. Example of regions of interest identification: green, pastures; cyan, water bodies; maroon, vineyards; magenta, permanently irrigated land; purple, broad-leaved forests

DOI: 10.1002/hyp

a pixel of the image, although clearly unrelated to any defined class, will have a small but not zero probability of belonging to the classes; therefore, it will be however assigned to one of them. To avoid coarse errors of classification, leading to overestimate the areas covered by the examined classes, one must admit that pixels having a probability $P(X|C_i)$ lower than a certain value b, called threshold value, cannot longer be considered as belonging to the class C_i with sufficient certainty, and so they have to be discarded. The threshold b can be assumed approximately equal to 5%. By increasing the value of b, the classification becomes more reliable, but with a consequent diminution of the classified areas.

The maximum likelihood algorithm used in this work was the one that gave the best results of this analysis, as it was able to extend the information drawn out by the regions of interest to the adjacent areas yielding the statistic relationships that allowed to discern the different elements and then 'decide' whether associate those pixels to the same typology of land cover identified.

The land use map obtained by processing the Landsat_7 ETM+ image of 26 September 1999 is shown in Figure 3.

A comparison immediately highlights the greater level of detail and precision in identifying the different elements in the thematic maps obtained by the supervised classification than the CORINE Land Cover (Figure 4).

Areas that resulted homogeneous in the CORINE have been distinguished in the regional land use maps, thanks to a greater punctual information. In fact, the applied methodology is able to distinguish the land use type within each pixel, even if isolated or included in areas characterized by a different type. An area of urban green, for example, can be easily distinguished from the surrounding infrastructures. Similar considerations can be made for the different types of agricultural land use. Moreover, the CORINE provides a temporal average of the territory representation, whereas the land use maps derived directly from Landsat images assure a temporal correspondence between ground and simulated reality.

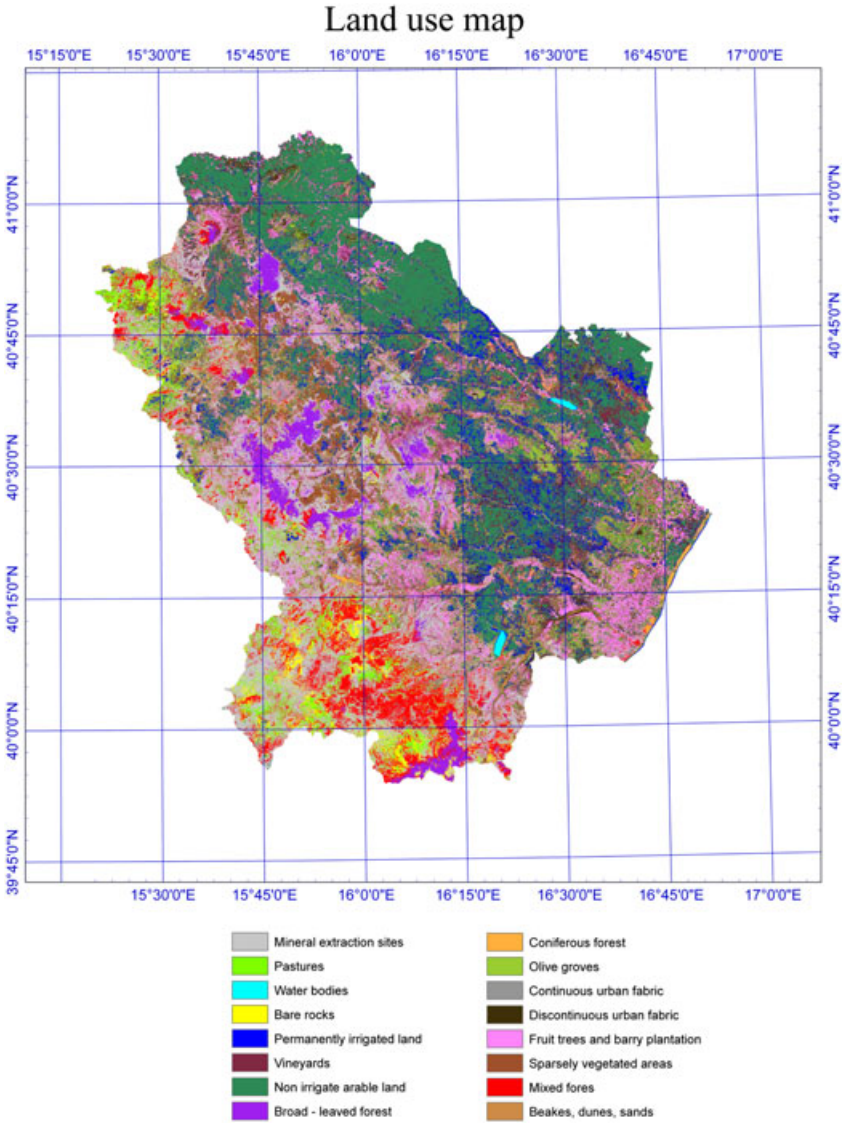

Figure 3. Land use thematic map of the Basilicata region (26 September 1999)

DOI: 10.1002/hyp

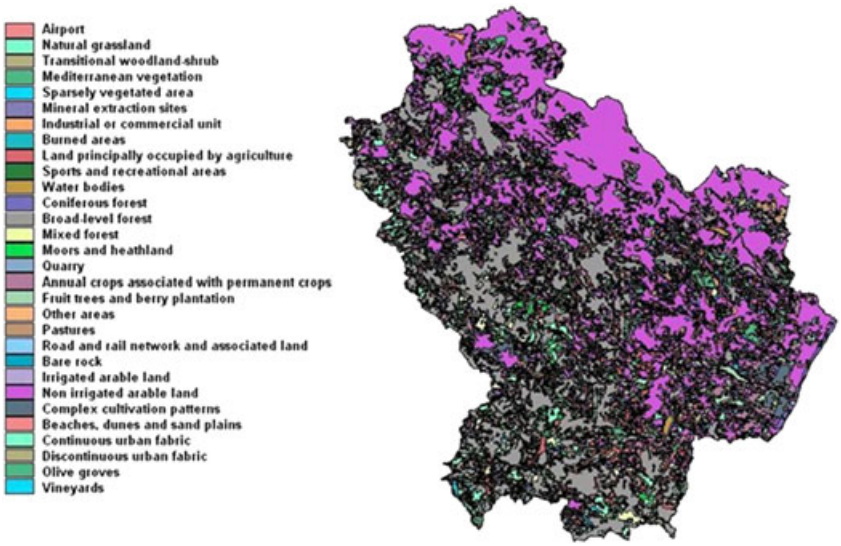

Figure 4. CORINE Land Cover 2000

The regional thematic maps of the Basilicata were used to evaluate the vegetation distribution, particularly the vegetation cover, P_v, vegetation height, h, and leaves dimension, s, required by the STSEB model, as described in the Application of STSEB Model section.

Classification accuracy analysis

The accuracy of the land use classification was evaluated by performing the analysis of the 'matrix of confusion' (Congalton *et al*., 1983; Story and Congalton, 1986; Congalton, 1988a, b), which compares the observed classes with the assigned category in the classified image. The classification accuracy assessment is a very delicate phase in thematic maps production from remote sensing as it provides an index of overall quality of the map, compares different classification algorithms and identifies possible errors in images processing (Hay, 1979).

Among the available validation procedures, the matrix of confusion is one of the most diffused because it not only provides more complete information than the simple percentage of pixels correctly classified but also characterizes the errors in the procedure allowing to improve the classification and to enrich the considerations that derive of it.

Several measures of accuracy derive from the matrix of confusion: the overall accuracy index, A, is one of the best known, which is obtained for each category of land use by dividing the number of pixels correctly classified by the total number of pixels of that class. However, many authors criticize the A method, asserting that some pixels may be randomly assigned to the correct class. They individualize in the agreement Khat of Cohen coefficient, K, the standard classification accuracy index (Congalton *et al*., 1983; Congalton, 1991).

The coefficient K is useful to express the overall degree of classification quality because it represents the difference between the achieved accuracy and what could be obtained from a completely random classification of the image: for example, a coefficient K equal to 0.8 means that the classifier did not commit the 80% of errors that a random classification would generated. The coefficient K ranges from 0 to 1; the more the coefficient tends to 1, the higher the agreement between real and classified data.

The values of accuracy obtained by processing the three Landsat images are reported in Table II. They show a good performance of the applied classification methodology and a good quality of the obtained maps, especially for the 26 September 1999 land use map ($A = 0.87$; $K = 0.78$).

The two vegetation parameters required by the STSEB model, as described in the Application of STSEB Model section, are leaves dimension, s, and vegetation height, h, that vary in space and time. Although the first can be considered a marginal parameter, the second has a major influence on the evapotranspiration retrieval, especially at a regional scale, where the vegetation distribution is sparse and varied.

Data of h and s are given in Tables III and IV, divided by type of vegetation. These values come from a combination of literature agronomic information (Pignatti, 1982) and a specific study on the characteristics of the Basilicata region typical vegetation in different seasons, performed by the authors.

Parameters h and s were distributed all over the region on the base of the Basilicata land use maps, so it was possible to combine them with remote sensing data.

APPLICATION OF STSEB MODEL

The STSEB model was applied to estimate the daily actual evapotranspiration on the Basilicata region,

Table II. Values of accuracy obtained for the pixel-oriented land use classification: overall accuracy index (A) and accuracy coefficient (K)

Landsat images date	A (overall accuracy)	K coefficient
26 September 1999	0.87	0.78
14 June 2002	0.68	0.64
26 May 2004	0.63	0.58

DOI: 10.1002/hyp

Table III. Values of vegetation height, h, and leaves size, s, assigned on the basis of the Basilicata thematic land use map (26 September 1999)

S_Value	Land use	h (m)	$s = 4A/P$ (m)
2.1.1.	Nonirrigated arable land	0.25	0.002
2.1.2.	Permanently irrigated land	0.25	0.002
2.2.1.	Vineyards	1	0.2
2.2.2.	Fruit trees and berry plantation	4	0.044
2.2.3.	Olive groves	4	0.017
2.3.1.	Pastures	0.2	0.014
3.1.1.	Broad-leaved forest	8	0.00196
3.1.2.	Coniferous forest	8	0.00196
3.1.3.	Mixed forest	12	0.091
3.3.3.	Sparsely vegetated areas	0.4	0.010

A, leaf area; P, leaf perimeter.

Table IV. Values of vegetation height, h, and leaves size, s, assigned on the basis of the Basilicata thematic land use maps (14 June 2002 and 26 May 2004)

S_Value	Land use	h (m)	$s = 4A/P$ (m)
2.1.1.	Nonirrigated arable land	0.5	0.027
2.1.2.	Permanently irrigated land	0.6	0.01
2.2.1.	Vineyards	1.5	0.075
2.2.2.	Fruit trees and berry plantation	3.0	0.044
2.2.3.	Olive groves	3	0.017
2.3.1.	Pastures	0.5	0.0384
3.1.1.	Broad-leaved forest	18	0.05
3.1.2.	Coniferous forest	10	0.014
3.1.3.	Mixed forest	12	0.091
3.3.3.	Sparsely vegetated areas	1.0	0.045

A, leaf area; P, leaf perimeter.

through the integration of field data, derived from agro-meteorological stations, and remote sensing data, coming from three Landsat satellite images.

The field data were used to estimate the typical parameters of the atmosphere evaporative power, such as wind speed, air temperature and global solar radiation. An interpolation of data measured at ALSIA stations allowed to draw up maps of air temperature, T_a, wind speed, u_c, and global solar radiation, R_t.

The remote sensing data were used for the evaluation of some 'critical' parameters of the evapotranspiration, such as the albedo, α, involved in the net radiation estimation, and the normalized difference vegetation index (NDVI), from which the vegetation cover, P_v, the surface emissivity, ε, and the radiometric land surface temperature, T_{rad}, depend.

Moreover, for the partition of the global sensible heat flux, H, in soil, H_s, and canopy, H_c, contributions, on the basis of partial areas, respectively, covered, the soil–vegetation–atmosphere scheme proposed by Sánchez *et al.* (2007, 2008a) was used.

The vegetation cover was characterized referring to the regional land use maps obtained by the application of the pixel-oriented supervised classification technique previously described (see Pixel-Oriented Land Use Classification section). They allowed to identify in detail not only the vegetated areas, distinguishing them from bare soil, but also the different types of vegetation, to which the specific height, h, and dimension of leaves, s, were associated, taking into account the reference seasons.

Atmospheric profiles derived from the radiosound of Brindisi (CLAR database), for each date, were introduced as input data in the MODTRAN 4.0 code. Visible and near-infrared bands were corrected, and an NDVI map was elaborated from each Landsat image. A procedure was applied to each NDVI map, based on the contrast in the bands 3 and 4, to identify the pixels completely covered by vegetation and those of bare soil. Emissivity maps were realized, assuming nominal values that suppose $\varepsilon_c = 0.985$ and $\varepsilon_s = 0.960$ for fully vegetated areas and bare soil, respectively.

The parameters involved in the simplified energy balance equation, such as net radiation, R_n, and sensible heat flux, H, were estimated and represented at a regional scale. The evapotranspiration was finally determined as a residual term of the energy balance equation, obtaining daily maps on the study area.

Methodology description

The applied methodology to estimate actual evapotranspiration is based on the STSEB equation, which models the soil–vegetation–atmosphere system:

$$R_n = G + H + LE \tag{2}$$

where R_n is the net radiation flux (W m^{-2}), G is the soil heat flux (W m^{-2}), H is the sensible heat flux (W m^{-2}) and LE is the latent heat flux (W m^{-2}).

This equation is the starting point of the models for estimating actual evapotranspiration using remote sensing techniques because most of them are based on the principle of energy conservation (Anderson *et al.*, 1997; Bastiaanssen *et al.*, 1998a, b; Su, 2001, 2002; Li *et al.*, 2005; Allen *et al.*, 2007a, b; etc.).

There are three surface energy balance approaches for estimating areal evaporation; the first one is the residual approach, which calculates R_n, G and H and then obtains the latent heat flux (LE) as the residual of the energy balance equation; the second approach uses a water stress index to estimate the relative evaporation; the third approach is to compute all components of the energy budget at the land surface with continuous land surface models, which include soil–vegetation–atmosphere transfer models (Kalma *et al.*, 2008).

According to Seguin and Itier (1983), for midday instantaneous and clear-day data:

$$\frac{H_d}{R_{nd}} = \frac{H_i}{R_{ni}} \tag{3}$$

where the subscripts i and d indicate, respectively, instantaneous and daily flux.

Moreover, because at a daily scale the soil heat flux, G, can be neglected (Seguin and Itier, 1983; Lagouarde and McAneney, 1992; Sánchez *et al.*, 2007), LE can be

DOI: 10.1002/hyp

obtained from Equation (2) and from instantaneous values of R_n and H corresponding to the satellite passing over the study area, using the coefficient R_{nd}/R_{ni}. Combining Equations (2) and (3):

$$LE_d = \frac{R_{nd}}{R_{ni}}(R_{ni} - H_i) \tag{4}$$

The instantaneous net radiation flux is evaluated through a balance equation between long-wave and short-wave radiations:

$$R_{ni} = (1 - \alpha)R_t + \varepsilon L_{sky} - \varepsilon\sigma T_{rad}^4 \tag{5}$$

where α (−) is the albedo, R_t (W m^{-2}) is the global short-wave radiative flux, ε (−) is the effective surface emissivity, L_{sky} (W m^{-2}) is the long-wave incoming radiation, T_{rad} (K) is the radiometric surface temperature (Sánchez *et al.*, 2008b) and σ is the Stefan–Boltzmann constant.

The surface albedo, α, was integrated by using the equation of Dubayah (1992), whereas a simple and operational equation proposed by Valor and Caselles (2005) was used to estimate the surface emissivity, ε.

In the STSEB approach proposed by Sánchez *et al.*, (2008a), the surface is divided in two components, canopy and surrounding bare soil. According to this configuration, the global sensible heat flux is a combination of soil and canopy contributions, H_s and H_c, proportionally to the vegetation cover, P_v, according to the following equation:

$$H = P_v H_c + (1 - P_v)H_s \tag{6}$$

where P_v was obtained through the Valor and Caselles (1996) equation, and H_s and H_c are expressed as follows:

$$H_c = \rho C_p \frac{T_c - T_a}{r_a^h} \tag{7a}$$

$$H_s = \rho C_p \frac{T_s - T_a}{r_a^a + r_a^s} \tag{7b}$$

In these equations, ρC_p is the volumetric heat capacity of the air (J K^{-1} m^{-3}), T_a (K) is the air temperature at the reference height, T_c and T_s (K) are canopy and soil radiometric temperatures, respectively, and r_a^h (sm^{-1}) is the aerodynamic resistance to the heat transfer between canopy and reference height, given by the following equation:

$$r_a^h = \frac{\left[\ln\left(\frac{z-d}{z_0}\right)\right]^2}{k^2 u} \tag{8}$$

in which it was set $d = (2/3)h$ and $z_0 = (1/10)h$, where h (m) is the height of vegetation, d (m) is the height of zero displacement, z_0 (m) is the surface roughness length, r_a^a (sm^{-1}) is the aerodynamic resistance to the heat transfer between the point $(z_0 + d)$ and the reference height (generally set equal to r_a^h), and r_a^s (sm^{-1}) is the aerodynamic resistance to heat flux in the boundary layer immediately above the soil surface, given by the Kustas and Norman (1999) equation:

$$r_a^s = \frac{1}{c(T_s - T_c)^{1/3} + bu_s} \tag{9}$$

where $c = 0.0025$ and $b = 0.012$ are two constants, and u_s (ms^{-1}) is the wind speed at the height above the ground surface where the effect of soil surface roughness on the free wind movement can be neglected. It is given by the following equation:

$$u_s = u_c \exp\left[-0.28\left(\frac{\mathrm{Ln}\left(\frac{1}{1-P_v}\right)}{0.5}\right)^{2/3} h^{1/3} s^{-1/3}\left(1 - \frac{0.05}{h}\right)\right] \tag{10}$$

where u_c (ms^{-1}) is the wind speed measured by the ALSIA agro-meteorological stations, and h and s (m) represent the height and leaves size of vegetation types in the study area (Tables III and IV).

RESULTS AND DISCUSSION

In this section, the actual evapotranspiration results obtained by applying the STSEB model to the study area are introduced. The comparison of these results with *in situ* measurements of evapotranspiration obtained by a weighing lysimeter located in the study area is also highlighted.

In Figures 5 and 6, the distributions on the Basilicata region Ionian band of vegetation height, h, and leaves size, s, respectively, corresponding to 26 September 1999 are represented. The map of instantaneous sensible heat flux, H, for the same date, obtained after evaluating canopy and soil radiometric temperatures, T_c and T_s, from the T_{rad} map, is represented in Figure 7.

At the time of Landsat passage, generally the surface radiometric temperature, T_{rad}, is higher than the air temperature, T_a, for the most places (H positive) (Moran *et al.*, 1997). It is possible that under particular surface conditions, T_a is slightly higher than T_{rad}, resulting in little negative values of H. Some errors in T_{rad}, T_c and T_s retrieval have also to be taken into account. Furthermore, the maps of air temperature and wind speed that are input data for sensible heat flux evaluation may contain some spatial uncertainties that would lead to improbable and excessive negative values of H. For this reason, areas with $H < 50$ W m^{-2} were not considered in the present analysis.

R_{nd}/R_{ni} values were calculated for each date from the ground-collected data. Because it was shown that the ratio R_{nd}/R_{ni} varies over time and with date and latitude of the site, but not with the vegetation type (Sobrino *et al.*, 2005; Sánchez *et al.*, 2007), a constant value was used for each image: 0.365 for the 1999 Landsat_7 image, 0.378 for the 2002 Landsat_7 image and 0.351 for the 2004 Landsat_5 image.

DOI: 10.1002/hyp

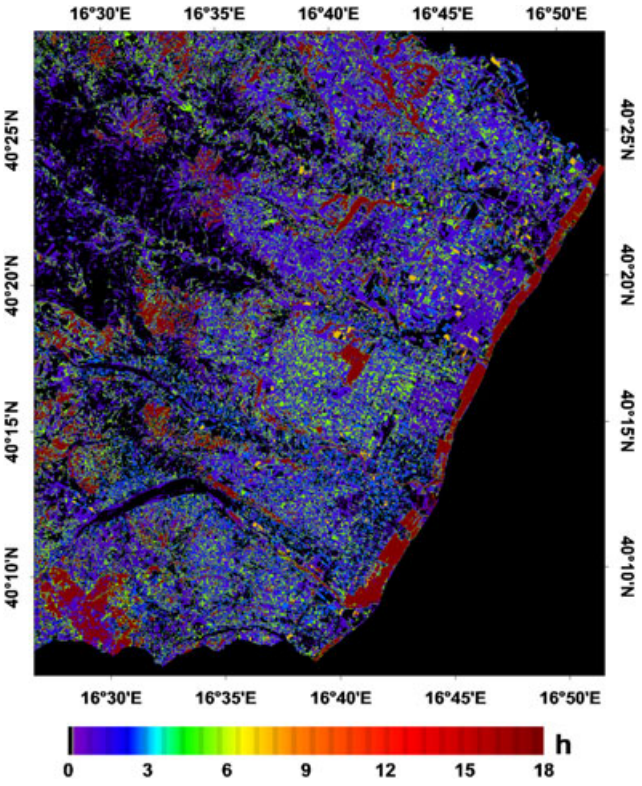

Figure 5. Distribution of the vegetation height, h (m), in the Basilicata Ionian band, South Italy (26 September 1999)

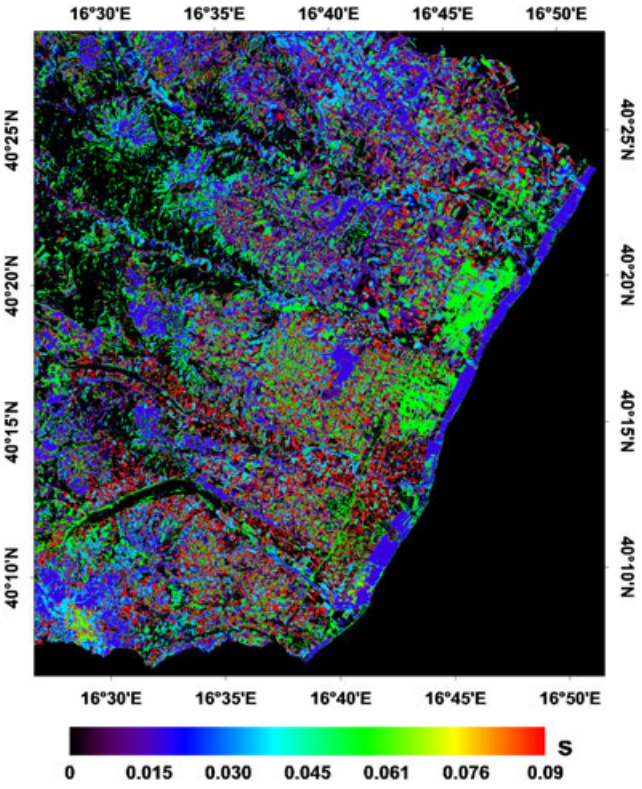

Figure 6. Distribution of the vegetation leaves dimension, s (m), in the Basilicata Ionian band, South Italy (26 September 1999)

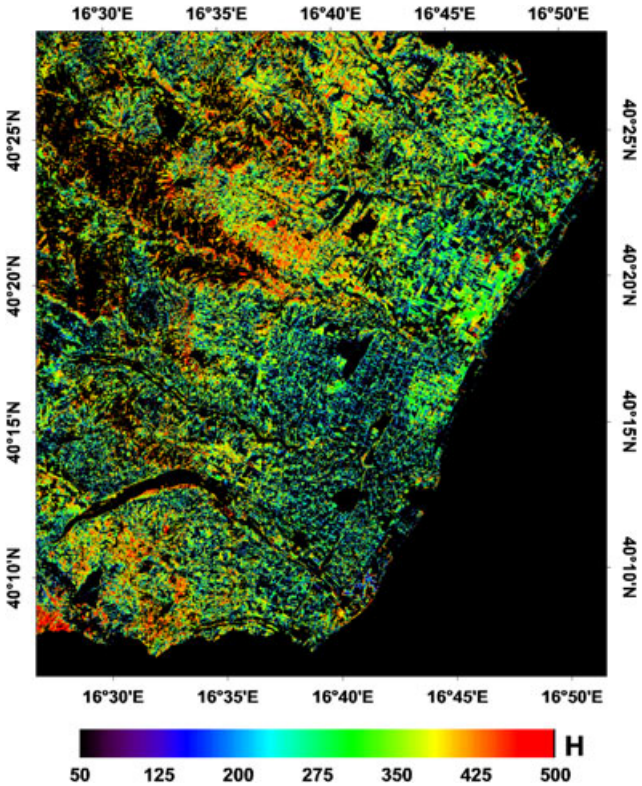

Figure 7. Map of instantaneous sensible heat flux, H (W m^{-2}), in the Basilicata Ionian band, South Italy (26 September 1999)

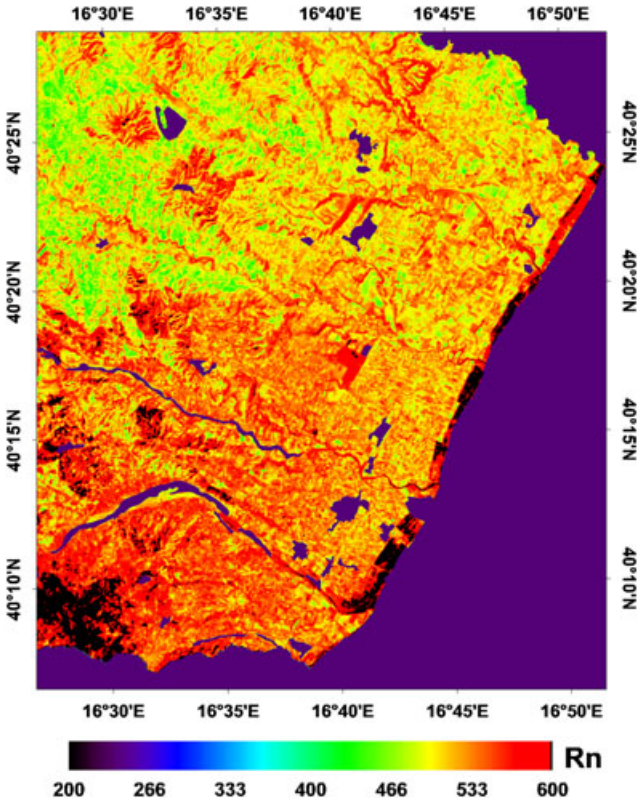

Figure 8. Map of instantaneous net radiation, R_n (W m^{-2}), in the Basilicata Ionian band, South Italy (26 September 1999)

In Figure 8, the map of instantaneous net radiation, R_n, estimated through Equation (5) for the 1999 Landsat_7 image is shown, whereas Figure 9 represents the daily actual evapotranspiration map, LE_d, for the same date. Finally, a frequency diagram of daily actual evapotranspiration values is represented in Figure 10.

DOI: 10.1002/hyp

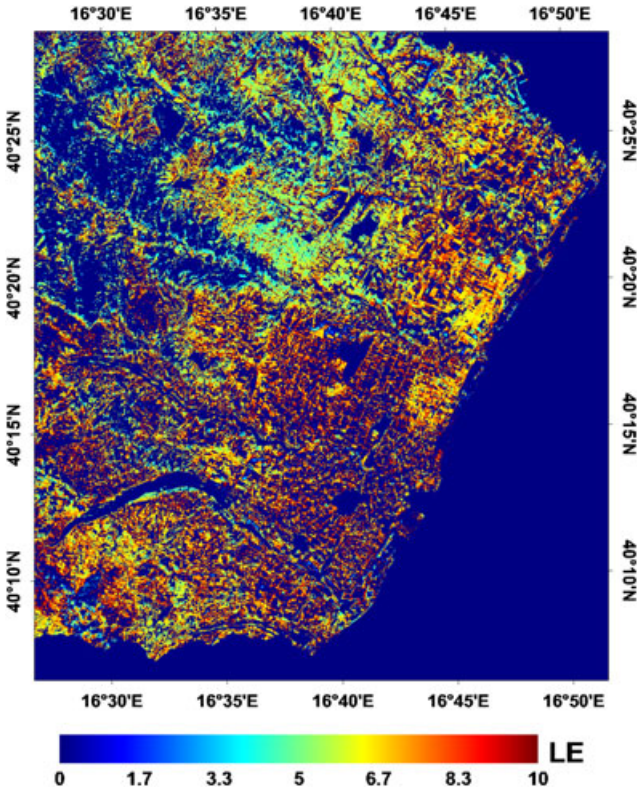

Figure 9. Map of daily actual evapotranspiration, LE_d (mm day^{-1}), in the Basilicata Ionian band, South Italy (26 September 1999)

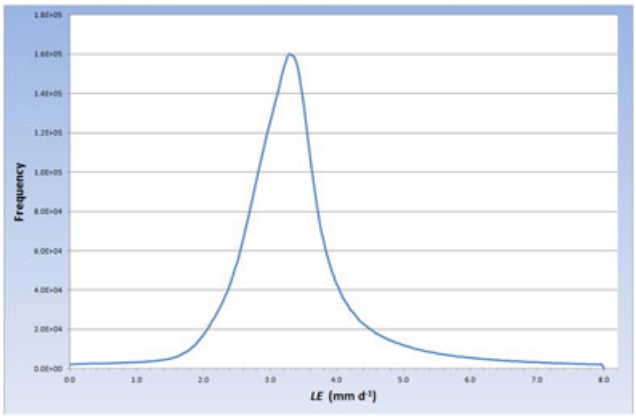

Figure 10. Frequency diagram of daily actual evapotranspiration values in the Basilicata Ionian band, South Italy (26 September 1999)

Table V shows the evapotranspiration results for the three dates, obtained using the regional land use maps, locally compared with the values measured by the weighing lysimeter at the experimental station 'E. Pantanelli' of Policoro (MT).

It can be noticed that the evapotranspiration results are very close to the measured values (each difference is less than 0.5 mm day^{-1}).

CONCLUSIONS

The aim of this work was to evaluate the applicability in a test area of a methodology for estimating actual evapotranspiration, based on a supervised land use classification within an STSEB model and on the integration of field and remote sensing data.

Table V. Results of estimated actual evapotranspiration compared with the values measured at the Policoro (MT) lysimetric station (Policoro, MT)

	LE_d (mm day^{-1}) estimated	LE_d (mm day^{-1}) measured	Differences ΔLE_d
26 September 1999 12:00 LT	2.0	1.9	0.1
14 June 2002 12:00 LT	2.8	2.9	−0.1
26 May 2004 12:00 LT	1.5	1.1	0.4

For such purpose, maps of land use specific of the study area (Basilicata region, South Italy) were realized by applying a pixel-oriented supervised classification technique to three Landsat TM and ETM+ images, taking advantage of the direct knowledge of the vegetation distribution in many sample areas. An accuracy analysis based on the overall accuracy index, A, and the agreement Khat of Cohen coefficient, K, showed a good performance of the applied classification method and a good quality of the obtained land use maps. Consequently, the aerodynamic resistance and sensible and latent heat fluxes in the soil–vegetation–atmosphere system, closely dependent on the vegetation distribution parameters (vegetation height, h, and leaves dimension, s), were also evaluated.

Maps of instantaneous sensible heat flux, H (W m^{-2}), instantaneous net radiation, R_n (W m^{-2}) and daily actual evapotranspiration, LE_d (mm day^{-1}), covering the Basilicata Ionian band were represented.

Finally, the results in terms of evapotranspiration were compared with the values measured at a lysimetric station located in Policoro (MT), in the Ionian zone of the Basilicata region.

The comparison showed that the estimated values of actual evapotranspiration are very close to those measured in the field. In particular, the differences between the estimated values and the measured ones are included within the range of acceptable errors for the specific variable (differences less than ± 0.5 mm day^{-1}), showing, in a first approximation, the applicability and validity of the proposed methodology.

However, it should be considered that these results need to be further validated through the processing of other satellite images because only one point of ground measurements is available for the validation. Furthermore, because the resulting values of evapotranspiration are very low (1.1–2.8 mm/day), it will be interesting to see how the model can reproduce the observed values in different seasons to prove the goodness of the model also in a not water limited period. A lysimetric measurements campaign is also planned for a next future, in the study area, to be able to further validate the obtained results in a great number of ground measurement sites.

DOI: 10.1002/hyp

ACKNOWLEDGEMENTS

The authors express their gratitude to ALSIA for providing agro-meteorological field data and Istituto di Scienze delle Produzioni Alimentari of Bari for providing lysimetric data. This work was funded by the Spanish Science and Innovation Ministry (Project CGL2010-17577) and Generalitat Valenciana (Project PROMETEO/2009/086).

REFERENCES

Allen RG, Tasumi M, Trezza R. 2007a. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC): model. *Journal of Irrigation and Drainage Engineering* **133**(4): 380–394. DOI:10.1061/(ASCE) 0733-9437(2007)133(4):(380).

Allen RG, Tasumi M, Trezza R. 2007b. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC): applications. *ASCE Journal of Irrigation and Drainage Engineering* **133**(4): 395–406.

Anderson MC, Norman, JM, Diak GR, Kustas WP, Mecikalski JR. 1997. A two-source time integrated model for estimating surface fluxes using thermal infrared remote sensing. *Remote Sensing of Environment* **60**: 195–216.

Bastiaanssen WGM, Menenti M, Feddes RA, Holtslag AAM. 1998a. A remote sensing surface energy balance algorithm for land (SEBAL). Part 1: Formulation. *Journal of Hydrology* **212–213**: 198–212.

Bastiaanssen WGM, Pelgrum H, Wang J, Ma Y, Moreno JF, Roerink GJ, van der Wal T. 1998b. A remote sensing surface energy balance algorithm for land (SEBAL). Part 2: Validation. *Journal of Hydrology* **212–213**: 213–229.

Congalton RG. 1988a. Using spatial autocorrelation analysis to explore errors in maps generated from remotely sensed data. *Photogrammetric Engineering and Remote Sensing* **54**: 587–592.

Congalton RG. 1988b. A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. *Photogrammetric Engineering and Remote Sensing* **54**: 593–600.

Congalton RG. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment* **37**: 35–46.

Congalton RG, Oderwald RG, Mead RA. 1983. Assessing Landsat classification accuracy using discrete multivariate statistical techniques. *Photogrammetric Engineering and Remote Sensing* **49**: 1671–1678.

Copertino VA, Scavone G, Telesca V, Caselles V, Sánchez JM, Valor E. 2006. Sviluppo di una procedura integrata di analisi per la stima dell'evapotraspirazione reale. *IDRA2006: Proceedings of XXX Convegno di Idraulica e Costruzioni Idrauliche*, Roma 10–15 Settembre 2006.

Dubayah R. 1992. Estimating net solar radiation using Landsat Thematic Mapper and digital elevation data. *Water Resource Research* **28**: 2469–2484.

Hay AM. 1979. Sampling designs to test land-use map accuracy. *Photogrammetric Engineering and Remote Sensing* **45**: 529–533.

Humes K, Hardy R, Kustas WP, Prueger J, Starks P. 2002. High spatial resolution mapping of surface energy balance components with remotely sensed data. *Thermal Remote Sensing in Land Surface Processes*. CRC Press: New York; 110–132.

Kalma JD, McVicar TR, McCabe MF. 2008. Estimating land surface evaporation: a review of methods using remotely sensed surface temperature data. *Surveys in Geophysics* **29**: 421–469.

Kustas WP, Norman JM. 1999. Evaluation of soil and vegetation heat flux predictions using a simple two-source model with radiometric temperatures for partial canopy cover. *Agricultural and Forest Meteorology* **94**: 13–29.

Lagouarde J-P, McAneney KJ. 1992. Daily sensible heat flux estimation from a single measurement of surface temperature and maximum air temperature. *Boundary-Layer Meteorology* **59**: 341–362.

Li F, Kustas WP, Prueger JH, Neale CMU, Jackson TJ. 2005. Utility of Remote Sensing based two-source energy balance model under low and high vegetation cover conditions. *Journal of Hydrometeorology* **6**(6): 878–891.

Moran MS, Humes KS, Pinter PJ. 1997. The scaling characteristics of remotely sensed variables for sparsely vegetated heterogeneous landscapes. *Journal of Hydrology* **190**: 337–362.

Pignatti S. 1982. Flora d'Italia. Edagricole–Bologna, Italy.

Sánchez JM, Caselles V, Niclòs R, Valor E, Coll C, Laurila T. 2007. Evaluation of the B-method for determining actual evapotranspiration in a boreal forest from MODIS data. *International Journal of Remote Sensing* **28**(5–6: 1231–1250.

Sánchez JM, Kustas WP, Caselles V, Anderson M. 2008a. Modelling surface energy fluxes over maize using a two-source patch model and radiometric soil and canopy temperature observations. *Remote Sensing of Environment* **112**(3): 1130–1143.

Sánchez JM, Scavone G, Caselles V, Valor E, Copertino VA, Telesca V. 2008b. Monitoring daily evapotranspiration at a regional scale from Landsat-TM and ETM+ data: application to the Basilicata region. *Journal of Hydrology* **351**(1–2): 58–70.

Sánchez JM, Caselles V, Kustas WP. 2008c. Estimating surface energy fluxes using a micro-meteorological model and satellite images. *Tethys, Journal of Mediterranean Meteorology & Climatology* **5**: 25–36.

Seguin B, Itier B. 1983. Using midday surface temperature to estimate daily evaporation from satellite thermal IR data. *International Journal of Remote Sensing* **4**(2): 371–383.

Sobrino JA, Gómez M, Jiménez-Muñoz JC, Olioso A, Chehbouni G. 2005. A simple algorithm to estimate evapotranspiration from DAIS data: application to the DAISEX campaigns. *Journal of Hydrology* **315**: 117–125.

Story M, Congalton RG. 1986. Accuracy assessment: a user's perspective. *Photogrammetric Engineering and Remote Sensing* **52**: 397–399.

Su Z. 2001. A Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes from point to continental scale. In *Advanced Earth Observation—Land Surface Climate*, Su Z, Jacobs C (eds). Publications of the National Remote Sensing Board (BCRS), Delft, The Netherlands, USP-2, 01–02, 184 pp.

Su Z. 2002. The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. *Hydrology and Earth System Sciences* **6**(1): 85–99 (HESS).

Valor E, Caselles V. 1996. Mapping Land Surface Emissivity from NDVI: application to European, African, and South American areas. *Remote Sensing of Environment* **57**: 167–184.

Valor E, Caselles V. 2005. Validation of the vegetation cover method for land surface emissivity estimation. *Recent Research Developments in Thermal Remote Sensing*. Research Signpost: Kerala (India). pp. 1–20.