

Comparison of multiple linear regression and artificial neural network models for downscaling TRMM precipitation products using MODIS data

D. D. Alexakis¹ · I. K. Tsanis^{1,2}

Received: 3 November 2015 / Accepted: 7 July 2016 / Published online: 14 July 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract Precipitation plays a significant role to energy exchange and material circulation in Earth's surface system. According to numerous studies, traditional point measurements based on rain gauge stations are unable to reflect the spatial variation of precipitation effectively. On the other hand, satellite remote sensing could solve this limitation by directly providing spatial distribution of rainfall over large areas. During the last years, the Tropical Rainfall Measuring Mission (TRMM) has provided researchers with a large volume of rainfall data used for the validation of atmospheric and climate models. However, due to its coarse resolution (0.25°) the improvement of its resolution appears as a fundamental task. The main aim of this study is to compare two different integrated downscaling-calibration approaches namely multiple linear regression analysis and artificial neural networks for downscaling TRMM 3B42 precipitation data. The statistical relationship among TRMM precipitation data and different environmental parameters such as vegetation, albedo, drought index and topography were tested in the island of Crete, Greece. Free distributed satellite data of coarse resolution such as those of MODIS sensor were incorporated in the overall analysis. Multiple linear regression as well as artificial neural network models was developed and applied, and extensive statistical analysis was performed by downscaling the TRMM products. The downscaled precipitation estimates as well as the TRMM

products were subsequently validated for their accuracy by using an independent precipitation dataset from a ground rain gauge network. The downscaling procedure succeeded to significant improvements of monthly precipitation estimation (100 % improvement in terms of spatial resolution) in terms of spatial analysis with means of satellite remote sensing.

Keywords Downscaling · TRMM · MODIS · Neural networks · Precipitation

Introduction

Precipitation has an essential impact on human activities (Michaelides et al. 2009). There are three main sources of precipitation estimates: rain gauge stations, ground radars and remote sensing technology. However, according to numerous studies traditional point measurements based on rain gauge stations cannot reflect the spatial variation of precipitation effectively. Furthermore, ground radar systems are generally aimed at monitoring of extreme events over limited time spans and are not suitable for long-term arrangements due to their limited range (Immerzeel et al. 2009). On the other hand, satellite remote sensing could potentially solve this limitation by directly providing spatial rainfall over large areas. There is a continuous concern for techniques that improve the accuracy of the interpretation of satellite imageries. This concern is reflected in international literature (Themistocleous et al. 2013). Satellite precipitation products could be successfully used as an alternative to sparse rain gauge networks. Furthermore, remote sensing techniques are ideal for the collection of continuous and repeated rainfall data throughout space and time (Curtarelli et al. 2014). However, applications of

✉ D. D. Alexakis
alexakis@hydromech.gr

¹ School of Environmental Engineering, Technical University of Crete, University Campus, Chania, Crete, Greece

² Department of Civil Engineering, McMaster University, Hamilton, ON, Canada

these products are still limited due to the lack of robust quality assessment (Tan et al. 2015).

One of the remote sensing satellites most widely used to retrieve rainfall is the Tropical Rainfall Measuring Mission (TRMM). During the last years, the Tropical Rainfall Measuring Mission (TRMM) has provided researchers with a large volume of rainfall data used for the validation of atmospheric and climate models. However, due to its coarse resolution (0.25°) the improvement of its resolution appears as a fundamental task. TRMM data have been used in various applications such as the applicability of rainfall estimates for distributed hydrological modeling over a flood-prone region (Li et al. 2009), correlating rainfall peaks and water discharges (Shaban 2009), stream flow forecasting (Su et al. 2008) and comparing ground rain gauge data with TRMM measurements (Shrivastava et al. 2014). In addition, several studies concern TRMM data correction/calibration (Dinku et al. 2007; Condom et al. 2011; Cheema and Bastiaanssen 2012; Mantas et al. 2014), data uncertainties and data downscaling (Duan and Bastiaanssen 2013).

According to Atkinson (2013), downscaling has an important role to play in satellite remote sensing. Products of finer spatial resolution are derived on either assumptions or prior knowledge about the character of the target spatial variation coupled with spatial optimization, spatial prediction through interpolation or direct information on the relation between spatial resolutions in the form of a regression model. In this context, reports from several studies have highlighted that the relationship between precipitation and other environmental factors, such as vegetation and topography, is variable at different scales. Therefore, there is a key issue in the process of downscaling: the best scale at which the relationship between precipitation and other environmental factors established can be used in the final downscaling algorithm (Jia et al. 2011).

Various downscaling approaches have been applied in different studies during the last years. In the vast majority of studies (Jia et al. 2011; Fang et al. 2013; Chen et al. 2014), the positive relation between vegetation and precipitation has been highlighted through the use of normalized difference vegetation index (NDVI). Specifically, it has been proved that 0.25° TRMM precipitation and 1-km NDVI have a statistical relationship and this relationship was used to acquire 1 km annual precipitation data. Other studies such as Fang et al. (2013) have incorporated more factors in their downscaling model such as temperature, humidity roughness and topographical aspect for downscaling the TRMM data. Furthermore, Ud Din et al. (2008) implemented weighted bilinear interpolation methodology and Geographically Weighted Regression (GWR) for downscaling TRMM 3B343 data (Chen et al. 2014).

The main objective of this study is to apply an integrated downscaling methodology based on environmental

information such as vegetation, topography, drought and albedo derived from MODIS satellite products. In this way the spatial downscaling attempts to capture the sub-grid heterogeneity while preserving the characteristics at the original scale (Fang et al. 2013). The study area is the island of Crete, located in the southeastern Mediterranean. Two different methodologies, namely multiple linear regression (MLR) and artificial neural network (ANN) analysis, were applied to downscale the TRMM 3B 342 precipitation fields from 25- to 1-km pixel spatial resolution. The final results were validated based on the observation of 20 rain gauge stations located all around the island of Crete.

Study area and data

Study area

The island of Crete occupies the southern part of Greece (Fig. 1). With an area of 8265 km^2 , Crete covers almost 6.3 % of the area of Greece. The mean elevation is 482 m ranging from sea level to 2450 m, and the average slope is 228 m/km with the topography fracturing into small catchments with ephemeral streams and karst geology. Crete has a typical Mediterranean island environment with about 53 % of the annual precipitation occurring in the winter, 23 % during autumn and 20 % during spring while there is negligible rainfall during summer (Koutroulis and Tsanis 2010). The average annual precipitation for a normal year in the island of Crete is approximately 934 mm (Tsanis and Naoum 2003). This fact in addition to non-uniform precipitation distribution in the island (a reduction of almost 300 mm from the western to the eastern part of the island and a strong orographic effect) makes the water availability a very small but crucial portion of the total supply (Tsanis et al. 2011) (Fig. 1). This was the main reason that within the context of this research the island was divided spatially in two separate parts (western and eastern) so as each part to be studied independently in terms of precipitation data downscaling.

Data

One of the most widely used remote sensing satellites to retrieve precipitation is the Tropical Rainfall Measuring Mission (TRMM). TRMM is the result of a partnership between National Aeronautics and Space Administration and Japan Aerospace Exploration Agency, with main objective of monitoring and studying the rainfall in tropical and subtropical regions (Kummerow et al. 1998). TRMM carries five sensors on its payload, a Precipitation Radar (PR), the TRMM Microwave Imager (TMI), a visible and infrared (IR) scanner, a cloud and earth radiant energy

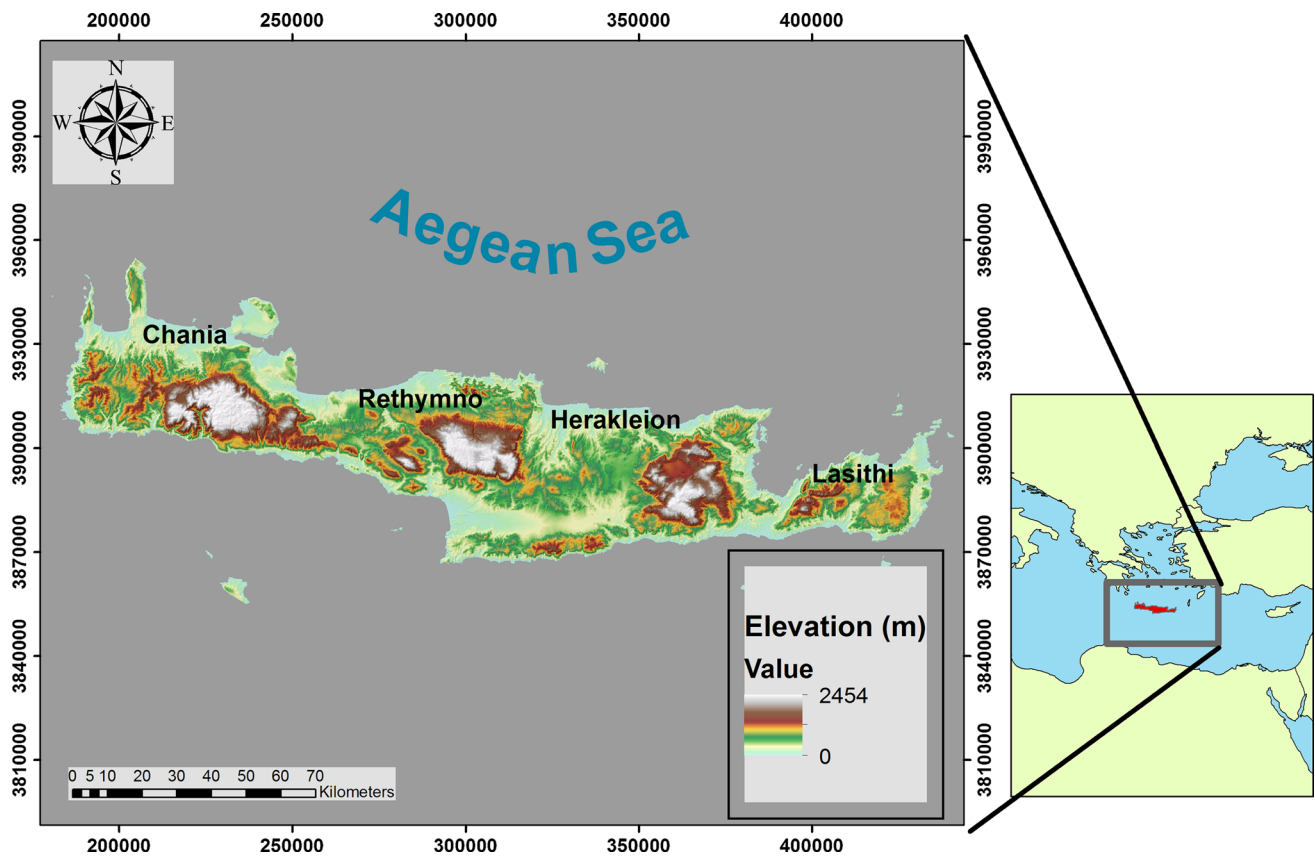


Fig. 1 Location and topography of the study area

sensor and a lightning imaging sensor. Each sensor has distinct purposes and measures energy at different ranges of the electromagnetic spectrum. The final digital product is based on an algorithm that merges microwave and infrared satellite estimates. However, the infrared satellite sampling is more frequent than the microwave sampling (AghaKouchak et al. 2009). For the needs of the study, TRMM 3B 342 images were used and analyzed for the period of 2012–2014. The specific data are available with a spatial resolution of $0.25^\circ \times 0.25^\circ$ and a temporal monthly variation, within 50° and 50° south global latitude.

MODIS NDVI data

NDVI data are widely used to detect vegetation regime with means of satellite remote sensing through the use of Near InfraRed (NIR) and Red spectral channel. Furthermore, spatial and temporal relationships between NDVI and precipitation indicate that there is generally a positive correlation between these two parameters (Chen et al. 2014). For the needs of the study, the Terra MODerate resolution Imaging Spectroradiometer (MODIS) monthly composite NDVI data of 1-km resolution (MOD13A3, collection v005) were collected from the Earth Observing

system data gateway. The data concerned the time period of 2012–2014.

Elevation data

TRMM precipitation data were analyzed in the context of topography. The digital elevation model (DEM) used for the needs of the study was derived from the digitization of topographical maps of the Hellenic Geographical Military Service. The spatial analysis of the extracted DEM was 20 m and depicted analytically the topography variation of the study area. Topography was incorporated in the overall analysis due to the fact that different precipitation rates can occur because of changes in air pressure, temperature, relative humidity and circulation of moist air, parameters that are affected by elevation.

Albedo data

Surface albedo, defined as the ratio of the total (hemispheric) reflected solar radiation flux to the incident flux upon the surface, quantifies the radiation interaction between the atmosphere and the land surface, and it can be directly related to precipitation regime in a local scale

(Wang et al. 2014). Albedo plays a crucial role in land surface climate and biosphere. The MODIS BRDF/albedo products (MCD43A) have been available since 2000 and provide high quality surface reflectance anisotropy retrievals over a variety of land surface type. For this study, MODIS albedo products (version 4) were analyzed covering the period of 2012–2014 at 1-km spatial and 16-day time resolution. The monthly albedo images were computed by averaging the two different albedo images acquired per month.

Drought index

Drought is a stochastic natural phenomenon that arises from considerable deficiency in precipitation. Drought indices are quantitative measures that characterize drought levels by assimilating data from one or several variables (indicators) such as precipitation and evapotranspiration into a single numerical value (Zargar et al. 2011). Monitoring drought can be achieved alternatively with the means of remote sensing precipitation products; in this case the remote sensing based drought indices synthesizing precipitation are developed and used in order to monitor the complex process of drought (Du et al. 2013). One of those indices is vegetation water supply index (VWSI) suggested by Cai et al. (2010) and is defined as:

$$VWSI = \frac{NDVI}{LST} \quad (1)$$

where LST is Land Surface Temperature MODIS product.

The MODIS LST is derived from two thermal (TIR) infrared band channels, namely 31 (10.78–11.28 μm) and 32 (11.77–12.27 μm). The product aims at retrieving LST with an error lower than 1 $^{\circ}\text{C}$ (± 0.7 $^{\circ}\text{C}$ standard deviation) in the range of -10 to 50 $^{\circ}\text{C}$ assuming the surface emissivity is known (Benali et al. 2012). In order to calculate the LST product in Celsius degrees, the following equation was used:

$$\text{Temperature } (^{\circ}\text{C}) = \text{DN} \times 0.02 - 273.15 \quad (2)$$

VWSI is showing the influence of drought on agriculture, and maps of summer drought over large areas. Using this method, vegetation growth can be closely monitored and the regional effects of summer drought can be recorded in detail. VWSI performs more efficiently on agriculture fields with densely covered vegetation areas. Both MODIS NDVI and LST datasets were utilized to develop VWSI for the island of Crete for the time period 2012–2014 (Fig. 2).

Meteorological data

A dataset with monthly precipitation data for the period 2012–2014 was used and incorporated in the study. The

data have been derived from the open precipitation database of the meteorological stations network, established in the island of Crete by the National Observatory of Athens (Table 1).

All the stations that are located above 1000 m were excluded from the analysis due to the orographic effect of precipitation and the errors that snow/wind causes in rain gauge data (Fig. 3).

Downscaling methodology

The TRMM data sets were downscaled on the basis of the assumption that there is a strong relationship between precipitation and other environmental factors such as vegetation, topography, albedo and drought. The TRMM downscaling was estimated through the equation below:

$$P_{\text{downscaled}} = P_{\text{estimated}} + P_{\text{residual}} \quad (3)$$

where $P_{\text{downscale}}$ is the downscaled precipitation, $P_{\text{estimated}}$ is the high spatial (1 km) resolution precipitation estimated from various environmental factors, P_{residual} is the residual between estimated precipitation and TRMM.

For the purposes of downscaling procedure, two different methodologies were used, namely multiple linear regression (MLR) analysis and artificial neural networks (ANNs). All data used in the study were of monthly time step and the overall analysis was mainly focused in the period of 2012–2014.

Rainfall independent areas

Four different “environmental” parameters such as vegetation, topography, albedo and drought were adopted to perform downscaling of TRMM data. Before starting the downscaling process, research for the rainfall dependency of the study areas was carried out. As the downscaling method is based on the relation between “environmental” parameters and precipitation, non-rainfall dependent areas such as urban areas and water bodies need to be excluded from the downscaling procedure. Thus, certain NDVI values such as those of water bodies are not related to precipitation and should therefore not be included in the NDVI-TRMM regression analysis (Duan and Bastiaanssen 2013). In order to exclude these areas, unsupervised classification (ISODATA algorithm) was performed on MODIS NDVI mean annual image for 2013 (Verlinde 2011). Class 2 of the final thematic map was assigned as the one that depicted the rainfall independent areas both for western and eastern part of Crete. Those areas were excluded from the downscaling procedure. Furthermore, inland water areas such as lakes and rivers were considered unsuitable for the downscaling procedure. This spatial

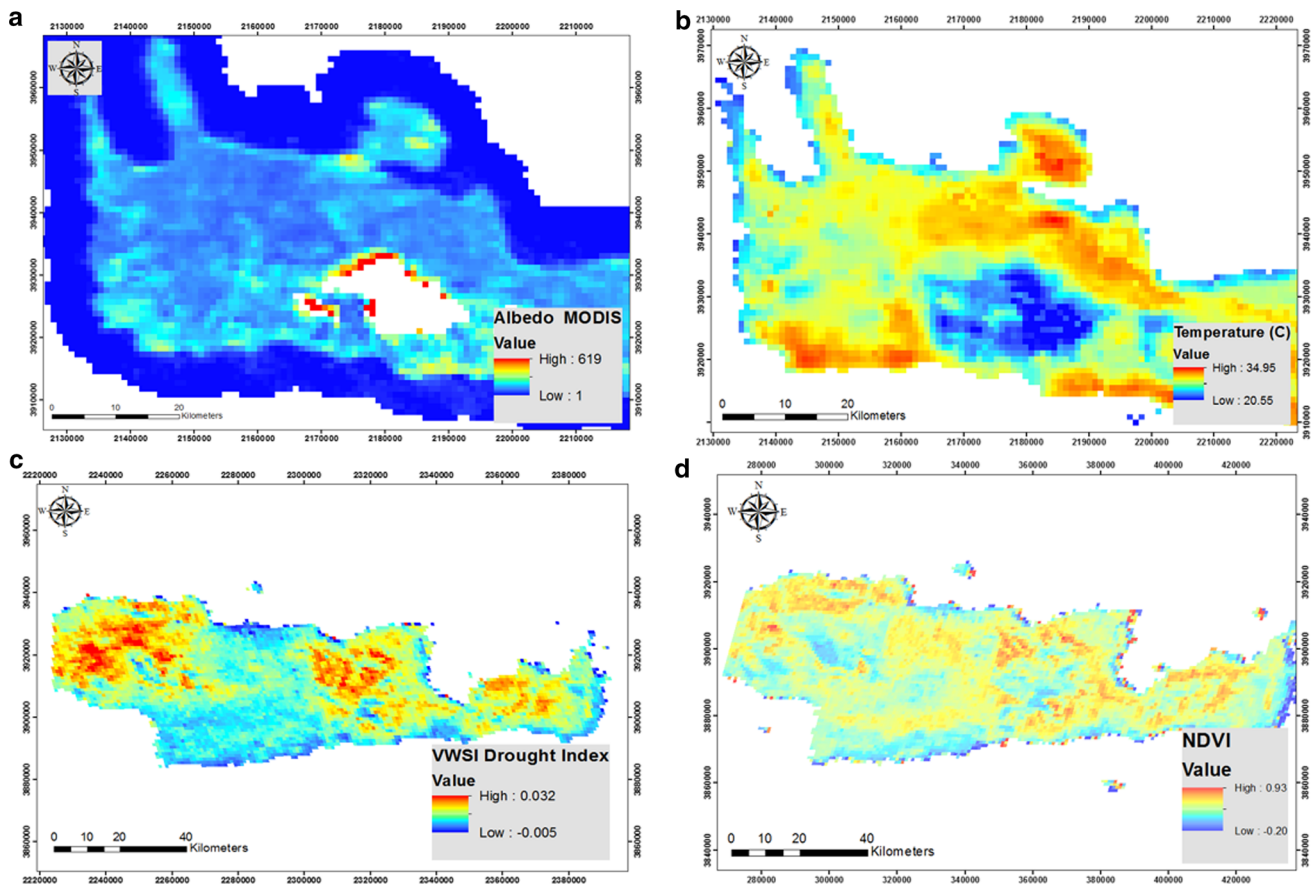


Fig. 2 a Albedo MODIS data (1 km); b LST MODIS product (Celsius); c VWSI Drought Index; d NDVI

information (inland water bodies) was derived from CORINE 2000 land use/land cover spatial database. Concerning the rivers, a buffer zone of 100 m was delineated around each segment and the total area was excluded from the overall downscaling procedure (Fig. 4). These outcomes were the reason to exclude from the overall analysis 13 rain gauge stations that are established on rainfall independent areas namely, “Chania Center,” “Chania,” “Fragma Potamon,” “Rethymno,” “Alikianos,” “Samaria Gorge,” “Siteia,” “Moires,” “Lentas,” “Metaxochori,” “Agios Nikolaos,” “Herakleion” and “Spili.”

Pre-processing statistical analysis

Before starting the downscaling process, Pearson correlation analysis was performed so as to check the least square fitting of TRMM precipitation data to ground precipitation measurements. The Pearson correlation analysis showed that the monthly areal rainfall TRMM 3B42 estimates are well correlated with that using the reference data (*r* values of 0.7 and 0.72 for western and eastern Crete accordingly). Furthermore, the *r*² of environmental parameters with ground monthly rainfall

measurements, ranged between 0.15 and 0.5 both for western and eastern Crete, indicating the positive correlation. However, it also justifies the implementation of multiple linear regression analysis in order to use combined information from all the different parameters and improve the overall accuracy.

Multiple linear regression analysis

The method of analysis used in multiple linear regression analysis is the method of least squares, which is simply a minimization of the sum of the squares of the deviations of the observed response from the fitted response (Naoum and Tsanis 2003). With precipitation being the dependent (response) variable, the model function involves both the predictor variables (NDVI, VWSI, albedo, elevation) and their corresponding parameters.

The general form of the final model is:

$$P = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 \tag{4}$$

where *P* is precipitation (mm month⁻¹), *x*₁ is NDVI, *x*₂ is albedo, *x*₃ is VWSI and *x*₄ is elevation and *b*₀, *b*₁, *b*₂, *b*₃, *b*₄ the corresponding parameters

Table 1 Rain gauge stations used in the study

	Name	Latitude	Longitude	Altitude (M)
1	Paleochora	35.21	23.66	17
2	Fragma Potamon	35.28	24.56	180
3	Rethymno	35.36	24.43	50
4	Samaria	35.30	23.91	1310
5	Spili	35.10	24.30	371
6	Falasarna	35.50	23.60	140
7	Fourfouras	35.21	24.71	472
8	Chania	35.53	24.06	131
9	Chania Center	35.50	24.00	5
10	Alikianos	35.45	23.91	63
11	Plakias	35.20	24.40	160
12	Ag. Nikolaos	35.18	25.7	14
13	Anogeia	35.28	24.88	807
14	Heraklion port	35.34	25.12	1
15	Ierapetra	35	25.4	87
16	Lentas	34.93	24.93	40
17	Metaxochori	35.13	25.14	424
18	Moires	35	24.8	146
19	Sitia	35.2	26.1	18
20	Tzermiado	35.2	25.3	336
21	Fourfouras	35.21	25.1	472
22	Heraklion West	35.3	25.1	83
23	Heraklion	35.31	25.15	90

Downscaling was performed as described below, following certain steps of analysis:

1. Aggregation of 1-km data of NDVI, VWSI, albedo and DEM to 0.25° by pixel average (resampling).
2. Application of multi-regression analysis and establishment of an empirical functional relationship between the (environmental parameters)^{0.25} and uncalibrated $P_{TRMM}^{0.25}$ for 2013. The parameters that passed the significance test (p value) were incorporated in the final model.
3. Estimation of monthly precipitation at 0.25° from the environmental parameters using the regression equation derived in step 2.
4. Preparation of a residual map at 0.25° by computing the difference between $P_{TRMM}^{0.25}$ and (environmental parameters)^{0.25} for every month of 2013. The residual map represents the amount of precipitation that cannot be explained by “Albedo.”
5. Interpolation of $P_{res}^{0.25}$ into a grid of 1-km pixels using spline interpolation methodology. This procedure was followed by other researchers as well (Immerzeel et al. 2009; Duan and Bastiaanssen 2013) due to the fact that the residual data are regular-spaced data and the spline interpolator is usually used for this kind of data.

6. Estimation of monthly precipitation at 1 km from (environmental parameters)^{1km} dataset using the regression equation derived in step 2.
7. Correction of the values of downscaled 1-km precipitation by adding the residual correction precipitation.

Neural networks analysis

Besides MLR, artificial neural networks (ANNs) were also applied to downscale precipitation data. The theory behind neural networks is based on an attempt to reproduce human learning processes (Aleotti and Chowdhury 1999). ANN has been used in different studies in the past to downscale coarse TRMM precipitation data (Kumar et al. 2007). It is an attractive and powerful numerical methodology to map complex relationships between different sets of observed variables (Tomassetti et al. 2009). An important advantage of the artificial neural network is its independency from the statistical distribution of the data and its ability to handle imprecise and fuzzy data (Conforti et al. 2014). An ANN consists of a collection of different neurons connected to each other. A connection occurs when the status of a neuron i is one of the inputs for another neuron j by means of a Weight (synapse) W_{ij} . The individual neurons are often called nodes of the network. The architecture of an ANN is defined by establishing how the individual neurons of the network are connected to each other. In order to estimate the number of the hidden layer nodes, equation proposed by Hecht–Nielsen (1987) was adopted:

$$N_g = 2 \times N_i + 1 \quad (5)$$

where N_g is the number of hidden nodes and N_i is the number of input nodes

A three-layer feed forward network consisting of an input layer (four neurons), one hidden layer (nine neurons) and one output layer was used as a network structure of 4–9–1. For the needs of the training procedure, 30,000 iterations were set as a threshold to terminate the procedure. The multi-layer perceptron (MLP) neural network was used for the application of ANN in the downscaling process. MLP consists of a set of layers, each of which is composed of a set of nodes and is trained with the back propagation algorithm using a set of examples of associated input and output values and at least one hidden layer. Tangent sigmoid function was used for transferring data from one layer to the other. The process of calibration, referred to as “training” of the ANN, consisted of the determination of all the weights (synapses) of the network based on the observed input/output patterns. During the ANN analysis, the consistency of training RMS with training testing values revealed the lack of overtraining and the accuracy of the process.

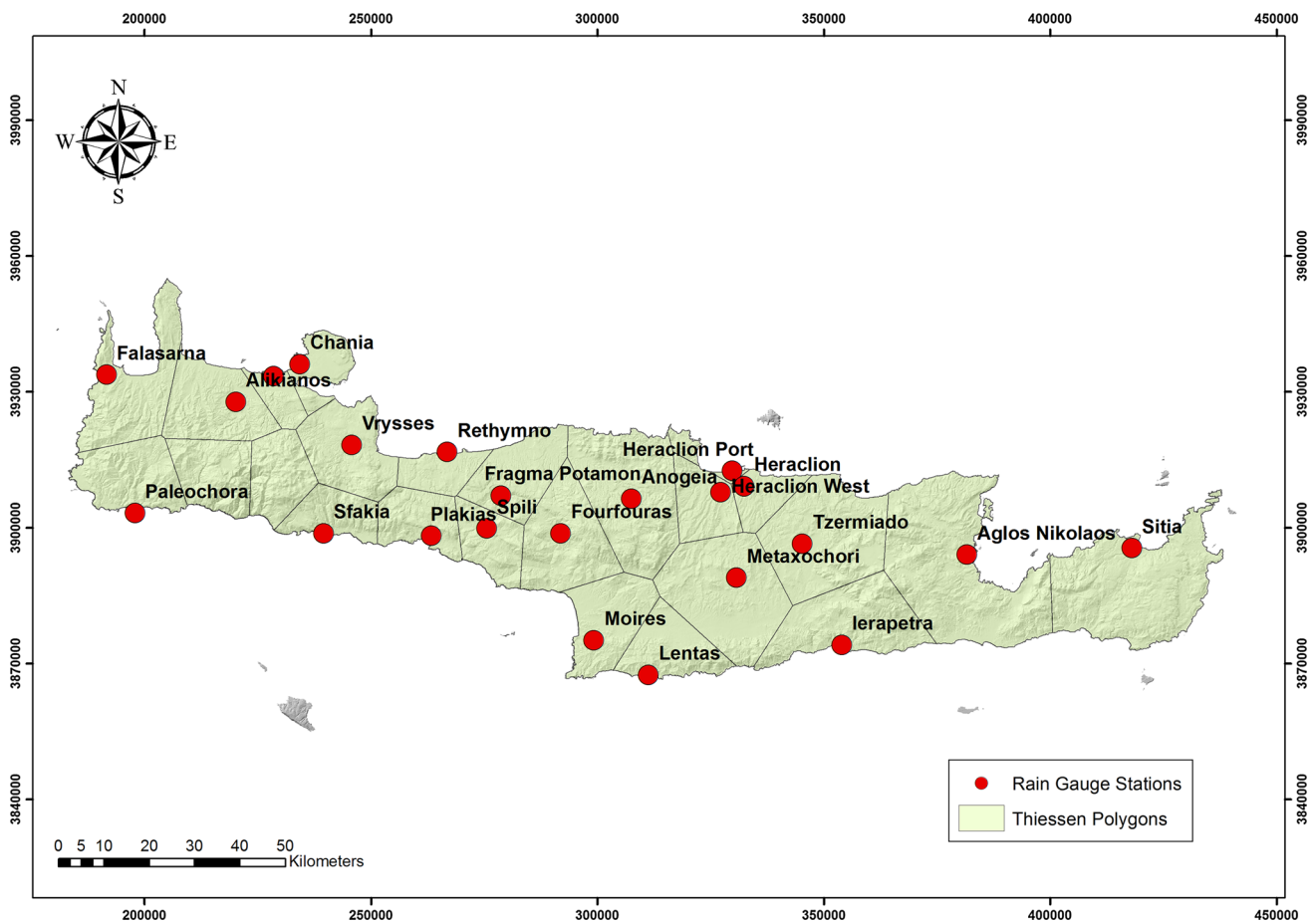


Fig. 3 Spatial distribution of meteorological stations and Thiessen polygon analysis

The steps used for downscaling in the case of ANN are almost identical to MLR analysis with some slight differences that are described below:

1. Aggregation of 1-km data of NDVI, VWSI, albedo and topography to 0.25° (environmental parameters) $^{0.25}$ by pixel average (resampling).
2. Use of ANN analysis for estimation of monthly precipitation ($P_{ANN}^{0.25}$) at 0.25° . Precipitation is estimated from monthly (environmental parameters) $^{0.25}$ data and monthly uncalibrated $P_{TRMM}^{0.25}$ data for every month of 2013.
3. Preparation of a residual map at 0.25° ($P_{res}^{0.25}$) by computing the difference between $P_{TRMM}^{0.25}$ and ($P_{ANN}^{0.25}$) for every month. The residual map represents the amount of precipitation that cannot be explained by the implementation of ANN.
4. Interpolation of ($P_{res}^{0.25}$) into a grid of 1-km pixels using spline interpolation methodology.
5. Estimation of monthly precipitation at 1 km from (environmental parameters) 1km dataset using the MLP

ANN algorithm. As a dependent image the map extracted from Step 6 is used.

6. Correction of the values of downscaled 1-km precipitation by adding the residual correction precipitation (Step 5).

Validation

The monthly ground precipitation data of 2013 were used to validate both the downscaling results and the TRMM products. In order to evaluate the downscaling methods quantitatively, two different statistical methods were calculated, the root-mean-square error (RMSE) and the bias according to:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - M_i)^2}{n}} \tag{6}$$

$$Bias = \frac{\sum_{i=1}^n P_i}{\sum_{i=1}^n M_i} - 1 \tag{7}$$

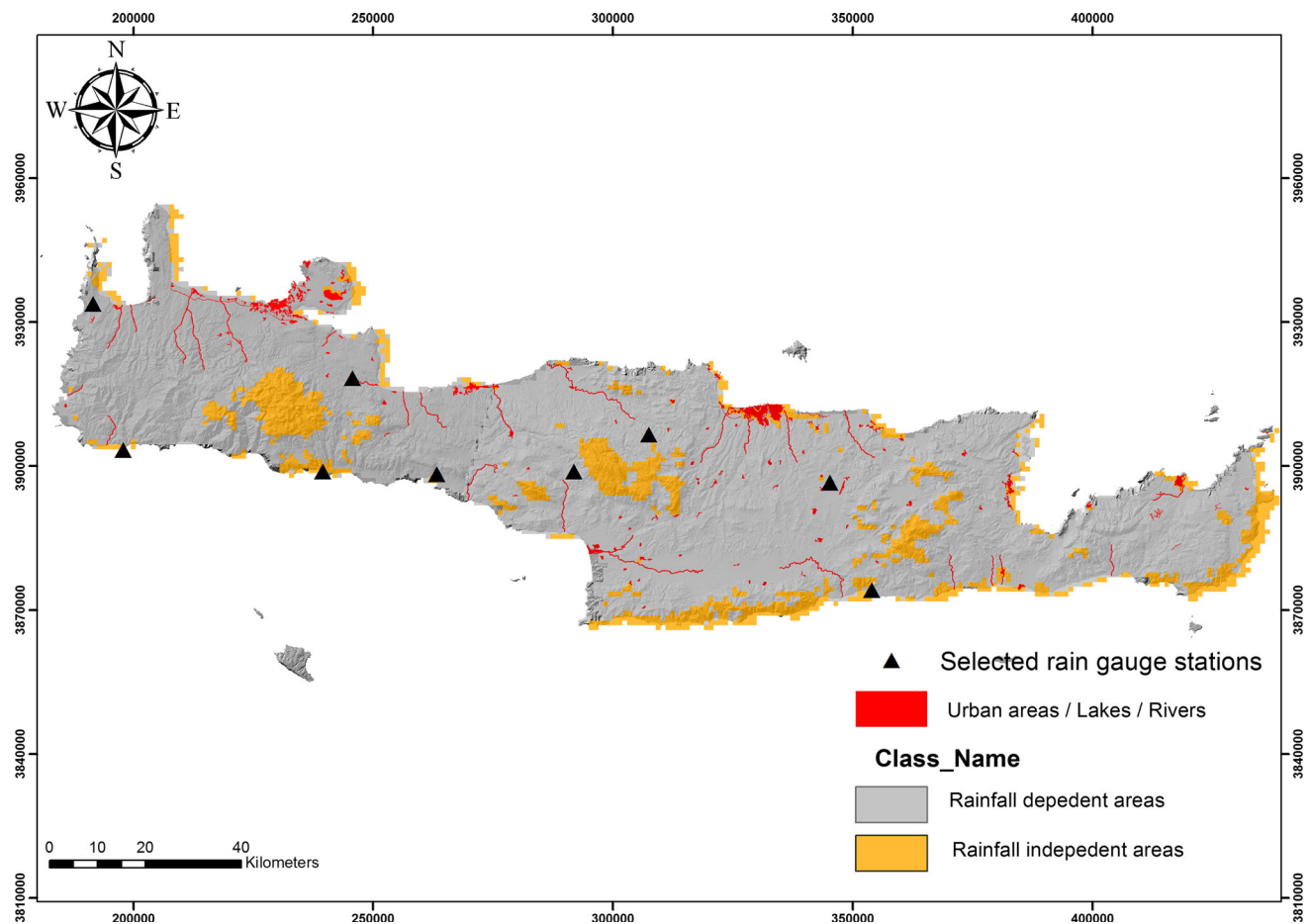


Fig. 4 Rainfall independent areas

P_i is the downscaled precipitation, M_i is the ground observed precipitation, n is the number of observations.

Results/discussion

After applying the validation process, multiple linear regression analysis and artificial neural networks were proved to have different performances in terms of accuracy. Both of these approaches were compared to TRMM precipitation values. The results denoted a better performance of MLR methodology in western Crete. In contrast, ANN outperformed MLR in eastern Crete.

Multiple linear regression analysis

Regarding western Crete MLR revealed a strong significant relationship ($p < 0.001$) between TRMM and albedo mainly due to vegetation cover regime. The p value for each term tests the null hypothesis that the coefficient is equal to zero (no effect). A low p value (< 0.05) indicates

that you can reject the null hypothesis. In case of eastern Crete, there is a correlation of TRMM with NDVI, albedo and VWSI. In both cases, MLR did not record any correlation of elevation with TRMM precipitation values. Furthermore, concerning model fitting, r^2 values for both cases ranged between 0.46 and 0.47. The results denoted that the derived regression relationship provides a medium description of the relationship between TRMM and the environmental parameters. The final regression equations for western and eastern Crete accordingly are:

$$P_i = 51.13 + 0.58 \times [\text{Albedo}] \quad (8)$$

$$P_i = 13.465 - 0.2179 \times [\text{Albedo}] + 340.37 \times [\text{VWSI}] + 90.73 \times [\text{NDVI}] \quad (9)$$

With the use of regression equations, the overall downscaling methodology was implemented and the results are presented in Fig. 5 (case of western Crete). As it was mentioned, the residual maps indicate areas where part of the precipitation cannot be explained by the environmental parameters only. Negative residual values depict areas where influence of some of the parameters is more than

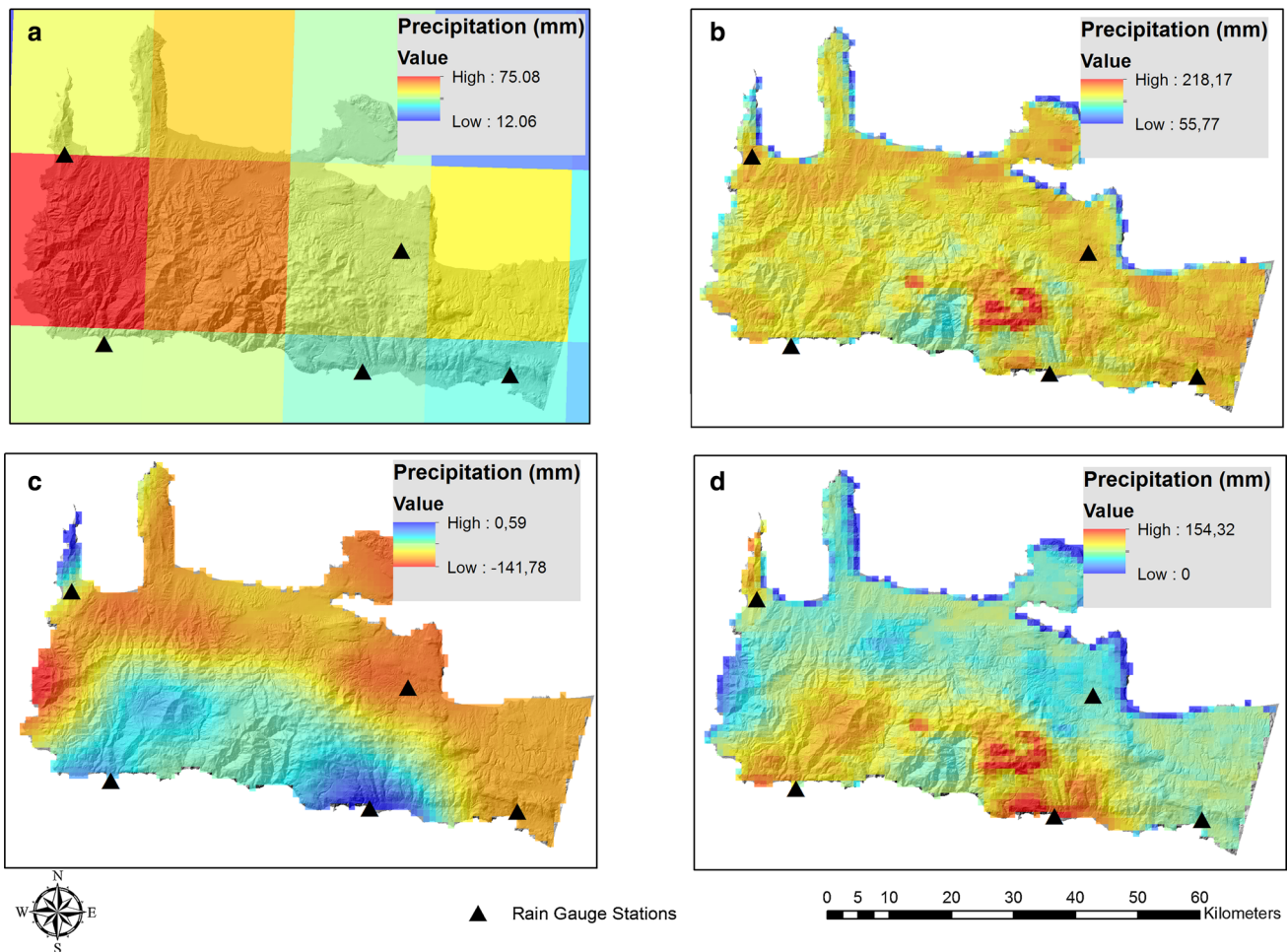


Fig. 5 **a** The calibrated TRMM 3B42 precipitation at 0.25° resolution; **b** the predictive precipitation at 1-km resolution; **c** the interpolated residuals at 1-km resolution; **d** the final downscaled result of precipitation at 1-km resolution

expected. In all these areas, an additional water source may be established (greener than expected), or some urban fabrics are established (greater values of albedo than expected) or are extremely drought due to a local severe drought incident such as surface salinity exposure.

As far as validation process is concerned, in case of western Crete, the RMSE values for MLR method ranged from 20.23 to 191.87 with optimum model performance during spring and summer period. Concerning bias, in few cases values were negative, indicating underestimation of the precipitation while precipitation during summer was generally overestimated (positive values) (Table 2). The fact that in some cases TRMM RMSE is almost identical to MLR RMSE is considered to be positive for the optimization of downscaled products in terms of spatial analysis. Concerning eastern Crete, RMSE values range from 11 to 182 and are slightly greater compared to those of TRMM and ANN. The general tendency in eastern Crete is an overestimation of precipitation.

Table 2 Statistics of validation results for eastern Crete

Month	Validation method					
	RMSE			BIAS		
	TRMM	MLR	ANN	TRMM	MLR	ANN
January	110.83	115.56	112.95	-0.43	-0.26	-0.35
February	51.29	59.04	56.55	0.50	0.57	-0.51
March	26.55	16.24	16.86	-0.21	1.39	0.56
April	36.16	29.93	49.07	-0.58	-0.28	-0.45
May	19.55	23.39	26.77	-0.19	-0.52	-0.30
June	26.59	39.95	14.36	-0.78	10.03	-0.28
July	3.72	35.06	2.36	-0.31	8.61	2.59
August	4.02	37.95	5.57	4.98	73.87	4.69
September	18.34	46.29	28.99	46.07	121.18	-0.29
October	25.78	60.75	59.19	-0.51	12.35	12.51
November	37.95	11.5	50.38	0.65	0.26	0.81
December	119.23	182.77	135.42	-0.36	-0.83	-0.40

Neural networks

The importance of an independent variable is a measure of how much the network's model predicted value changes for different values of independent variable. A sensitivity analysis to compute the importance of each predictor was applied. The final chart shows that the results for western Crete are mostly dominated by albedo and NDVI, whereas they are very slightly affected by elevation parameter, as in the case of MLR. Similarly, albedo is the most important and elevation is the less important parameter in eastern Crete (Fig. 6).

The ANN model that was used to downscale TRMM seems to perform reasonably well with mean r square values of 0.65 and 0.55 for eastern and western Crete, respectively (Fig. 7). In contrast to traditional statistical methods, ANN models provide dynamic output as further data are fed to them, while they do not require performing and analyzing sophisticated data methodologies (Kitikidou and Iliadis 2012). Concerning eastern Crete, ANN outperforms MLR methodology especially during summer period (Table 2). Regarding bias validation methodology, most of ANN values are negative meaning that precipitation is generally underestimated (Fig. 8). In western Crete, ANN performance is less successful in terms of RMSE and the model has optimal performance only during summer period (Table 3).

Overall analysis

The overall results denote that most accurate predictions are for months with moderate or low amounts of precipitation. This suggests that the ANNs have learned

partners that are generally common in the data, but have not learned partners that deviate from the mean monthly precipitation. In summary, concerning eastern Crete, ANNs slightly better performance is mainly related to the general high values of albedo due to extensive desertification phenomena occurring in that area. In this case, albedo succeeds to describe perfectly the vegetation regime of the study area. Hence, the high importance of albedo in ANNs model performance affects positively the overall result. Regarding MLR, the incorporation of all the environmental parameters in the overall equation impacts negatively the models accuracy. On the other hand, the performance of MLR in western Crete in terms of RMSE is slightly better compared to ANNs model mainly due to the incorporation of only albedo parameter in the final equation. Furthermore, the relatively low values of albedo parameter in western Crete have a negative impact on ANNs model performance that is generally highly affected by albedo.

Conclusions

Precipitation is a major factor in the entire hydrological process which greatly influences runoff generation. Its accurate estimation is a key for improving hydrological simulations and forecasting natural hazards such as floods. The purpose of this study was to compare two different integrated downscaling-calibration approaches namely multiple linear regression analysis and artificial neural networks for downscaling TRMM 3B42 precipitation data in order to improve monthly precipitation estimates resolution from 25×25 km to 1×1 km. Furthermore, the

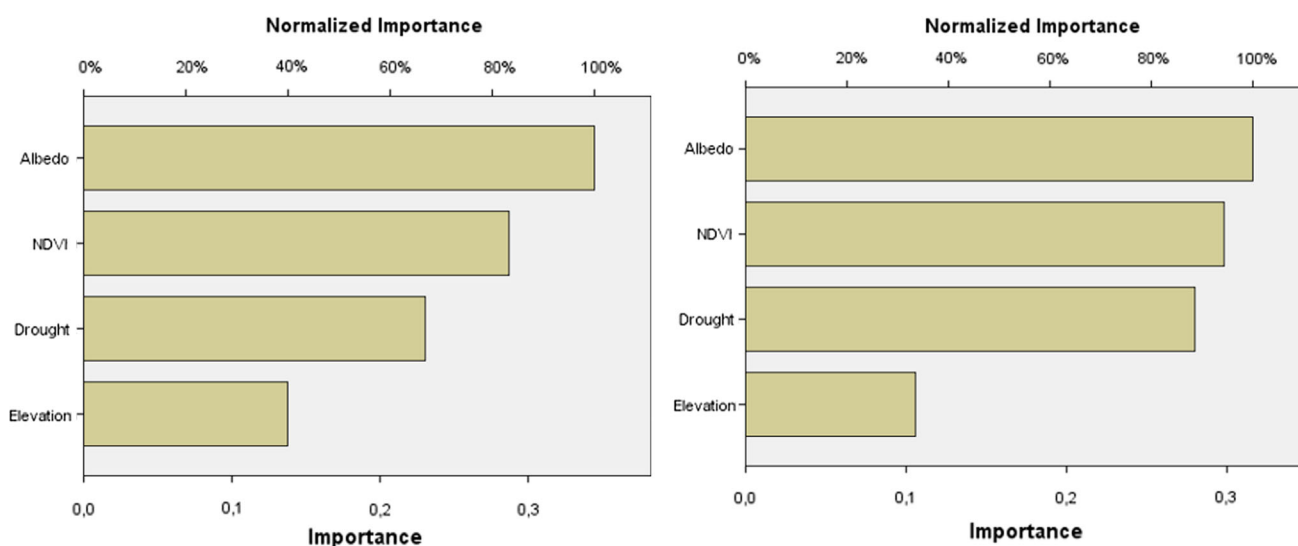


Fig. 6 a Neural network multi-layer perceptron independent variable important chart for western Crete; b eastern Crete

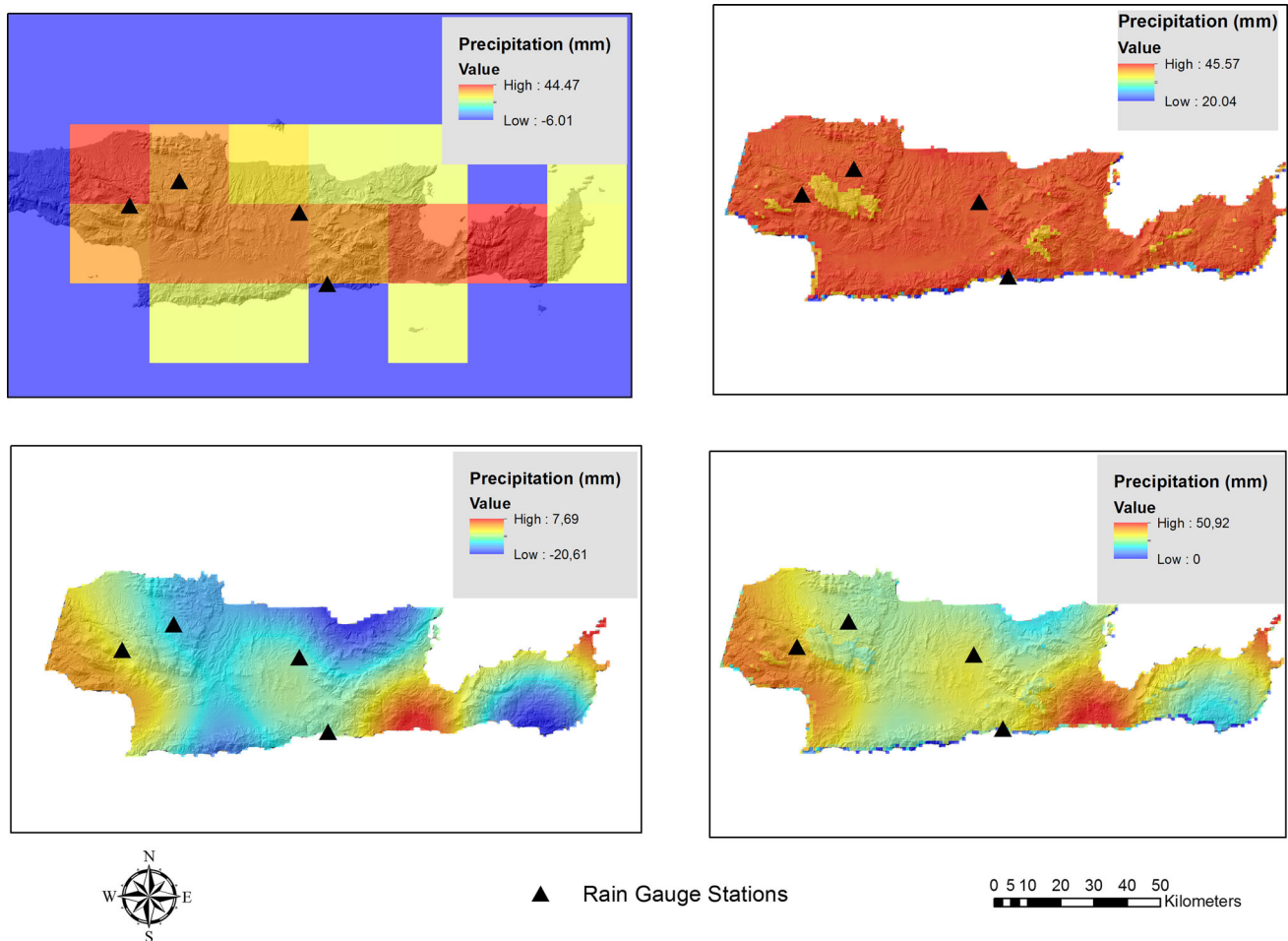


Fig. 7 **a** The calibrated TRMM 3B42 precipitation at 0.25° resolution; **b** the predictive precipitation at 1-km resolution; **c** the interpolated residuals at 1-km resolution, **d** the final downscaled result of precipitation at 1-km resolution

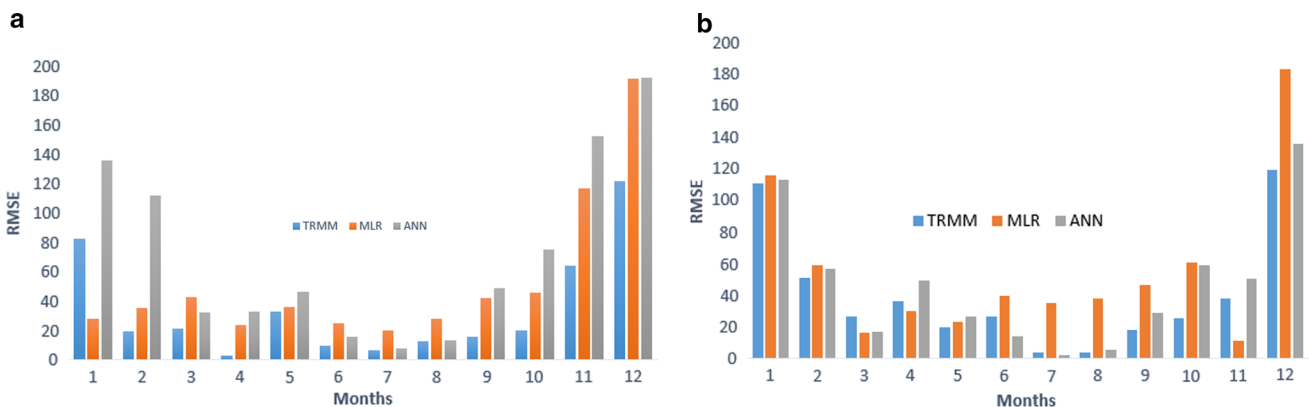


Fig. 8 RMSE comparative analysis of TRMM, MLR and ANN methodologies with the ground precipitation measurements for 2013. **a** Western Crete; **b** eastern Crete

exclusive use of MODIS data in the downscaling process pointed out the potential of free distributed satellite data of coarse resolution in satellite image processing procedure. The study area was the island of Crete in southeastern

Mediterranean and the main input was MODIS NDVI, LST and albedo data as well as regional topography.

The proposed downscaling methods assume that the relationship between precipitation and other environmental

Table 3 Statistics of validation results for western Crete

Month	Validation method					
	RMSE			Bias		
	TRMM	MLR	ANN	TRMM	MLR	ANN
January	82.37	27.83	135.88	-0.27	0.12	-0.37
February	19.76	35.37	112.17	0.127	0.38	0.16
March	21.54	42.80	32.12	1.065	1.63	0.42
April	3.344	24.01	33.17	0.09	-0.27	-0.36
May	33.02	35.99	46.48	1.2	1.79	1.08
June	9.846	24.97	15.64	5.16	6.39	11.52
July	6.55	20.23	7.80	18.98	30.95	12.57
August	12.72	27.84	13.57	-1	-1	-1
September	15.65	41.92	48.98	1.51	3.21	5.65
October	20.30	45.68	75.18	3.52	7.34	0.21
November	64.025	116.9634	152.26	0.486	0.9296	0.6889
December	121.91	191.8756	192.12	-0.046	-0.00085	-0.0005

variables varies spatially, but is the same in a local region. The main contribution of our research is the development of two integrated downscaling methodologies, the improvement of precipitation data spatial resolution, the overall comparative analysis of the different methodologies and the incorporation of free distributed satellite MODIS data in the overall process. Concerning research results, both techniques improved considerably the accuracy of precipitation record in terms of spatial resolution compared to TRMM coarse resolution. These approaches are characterized by low cost and if improved can replace ground instrumental techniques. The extracted 1-km resolution is suitable for hydrological modeling in catchment areas. However, the NDVI-based downscaling methods are applicable only to land surfaces and cannot be implemented to water bodies and urban areas (negative NDVI values). MLR methodology performed slightly better in western Crete while on the other hand ANN outperformed MLR in the eastern part. This result is mainly due to albedo parameter that deviates between the two study areas and highly affects MLR and ANNs models performance. Furthermore, it highlighted the potential of ANN to downscale successfully TRMM products using more than one variable. The research of the relationships between precipitation and related environmental factors revealed that in case of island of Crete environmental parameters such as albedo, drought index and NDVI can interpret reasonably well both precipitation variation and distribution. It was proved that there is no clear relationship between precipitation and topography. Furthermore, it was highlighted that the two different methodologies can be implemented supplementary in cases of downscaling approaches. The overall approach is generic in nature and can be used as a

road map for future improved downscaling techniques that could be developed for better spatial and temporal scales.

Acknowledgments The research reported in this paper was fully supported by the “ARISTEIA II” Action (“REINFORCE”/General Secretariat for Research and Technology, Hellas) of the “Operational Education and Life Long Learning programme” and was co-funded by the European Social Fund (ESF) and National Resources.

References

- AghaKouchak A, Nasrollahi N, Habib E (2009) Accounting for uncertainties of the TRMM satellite estimates. *Remote Sens* 1:606–619. doi:[10.3390/rs1030606](https://doi.org/10.3390/rs1030606)
- Aleotti P, Chowdhury R (1999) Landslide hazard assessment: summary review and new perspectives. *Bull Eng Geol Environ* 58:21–44. doi:[10.1007/s100640050066](https://doi.org/10.1007/s100640050066)
- Atkinson PM (2013) Downscaling in remote sensing. *Int J Appl Earth Obs Geoinf* 22:106–114. doi:[10.1016/j.jag.2012.04.012](https://doi.org/10.1016/j.jag.2012.04.012)
- Benali A, Carvalho AC, Nunes JP et al (2012) Estimating air surface temperature in Portugal using MODIS LST data. *Remote Sens Environ* 124:108–121. doi:[10.1016/j.rse.2012.04.024](https://doi.org/10.1016/j.rse.2012.04.024)
- Cai G, Du M, Liu Y (2010) Regional drought monitoring and analysing using MODIS data—a case study in Yunnan Province. In: *Computer and computing technologies in agriculture IV. IFIP Advances in information and communication technology*, vol 345. pp 243–251. doi:[10.1007/978-3-642-18336-2_29](https://doi.org/10.1007/978-3-642-18336-2_29)
- Cheema MJM, Bastiaanssen WGM (2012) Local calibration of remotely sensed rainfall from the TRMM satellite for different periods and spatial scales in the Indus Basin. *Int J Remote Sens* 33:2603–2627. doi:[10.1080/01431161.2011.617397](https://doi.org/10.1080/01431161.2011.617397)
- Chen F, Liu Y, Liu Q, Li X (2014) Spatial downscaling of TRMM 3B43 precipitation considering spatial heterogeneity. *Int J Remote Sens*. doi:[10.1080/01431161.2014.902550](https://doi.org/10.1080/01431161.2014.902550)
- Condom T, Rau P, Espinoza JC (2011) Correction of TRMM 3B43 monthly precipitation data over the mountainous areas of Peru during the period 1998–2007. *Hydrol Process* 25:1924–1933. doi:[10.1002/hyp.7949](https://doi.org/10.1002/hyp.7949)

- Conforti M, Pascale S, Robustelli G, Sdao F (2014) Evaluation of prediction capability of the artificial neural networks for mapping landslide susceptibility in the Turbolo River catchment (Northern Calabria, Italy). *Catena* 113:236–250. doi:[10.1016/j.catena.2013.08.006](https://doi.org/10.1016/j.catena.2013.08.006)
- Curtarelli MP, Rennó CD, Alcântara EH (2014) Evaluation of the tropical rainfall measuring mission 3B43 product over an inland area in Brazil and the effects of satellite boost on rainfall estimates. *J Appl Remote Sens* 8:083589. doi:[10.1117/1.JRS.8.083589](https://doi.org/10.1117/1.JRS.8.083589)
- Dinku T, Ceccato P, Grover-Kopec E et al (2007) Validation of satellite rainfall products over East Africa's complex topography. *Int J Remote Sens* 28:1503–1526. doi:[10.1080/01431160600954688](https://doi.org/10.1080/01431160600954688)
- Du L, Tian Q, Yu T et al (2013) A comprehensive drought monitoring method integrating MODIS and TRMM data. *Int J Appl Earth Obs Geoinf* 23:245–253. doi:[10.1016/j.jag.2012.09.010](https://doi.org/10.1016/j.jag.2012.09.010)
- Duan Z, Bastiaanssen WGM (2013) First results from Version 7 TRMM 3B43 precipitation product in combination with a new downscaling-calibration procedure. *Remote Sens Environ* 131:1–13. doi:[10.1016/j.rse.2012.12.002](https://doi.org/10.1016/j.rse.2012.12.002)
- Fang J, Du J, Xu W et al (2013) Spatial downscaling of TRMM precipitation data based on the orographical effect and meteorological conditions in a mountainous area. *Adv Water Resour* 61:42–50. doi:[10.1016/j.advwatres.2013.08.011](https://doi.org/10.1016/j.advwatres.2013.08.011)
- Immerzeel WW, Rutten MM, Droogers P (2009) Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula. *Remote Sens Environ* 113:362–370. doi:[10.1016/j.rse.2008.10.004](https://doi.org/10.1016/j.rse.2008.10.004)
- Jia S, Zhu W, Lu A, Yan T (2011) A statistical spatial downscaling algorithm of TRMM precipitation based on NDVI and DEM in the Qaidam Basin of China. *Remote Sens Environ* 115:3069–3079. doi:[10.1016/j.rse.2011.06.009](https://doi.org/10.1016/j.rse.2011.06.009)
- Kitikidou K, Iliadis L (2012) Developing neural networks to investigate relationships between air quality and quality of life indicators. In: *Air pollution-monitoring, modelling and health*, vol 1. pp 245–258. doi:[10.5772/34609.s](https://doi.org/10.5772/34609.s)
- Koutroulis AG, Tsanis IK (2010) A method for estimating flash flood peak discharge in a poorly gauged basin: case study for the 13-14 January 1994 flood, Giofiros basin, Crete, Greece. *J Hydrol* 385:150–164. doi:[10.1016/j.jhydrol.2010.02.012](https://doi.org/10.1016/j.jhydrol.2010.02.012)
- Kumar R, Das IML, Gairola RM et al (2007) Rainfall retrieval from TRMM radiometric channels using artificial neural networks. *Indian J Radio Space Phys* 36:114–127
- Kummerow C, Barnes W, Kozu T et al (1998) The tropical rainfall measuring mission (TRMM) sensor package. *J Atmos Ocean Technol* 15:809–817. doi:[10.1175/1520-0426\(1998\)015<0809:TTRMMT>2.0.CO;2](https://doi.org/10.1175/1520-0426(1998)015<0809:TTRMMT>2.0.CO;2)
- Li L, Hong Y, Wang J et al (2009) Evaluation of the real-time TRMM-based multi-satellite precipitation analysis for an operational flood prediction system in Nzoia Basin, Lake Victoria, Africa. *Nat Hazards* 50:109–123. doi:[10.1007/s11069-008-9324-5](https://doi.org/10.1007/s11069-008-9324-5)
- Mantas VM, Liu Z, Caro C, Pereira AJSC (2014) Validation of TRMM multi-satellite precipitation analysis (TMPA) products in the Peruvian Andes. *Atmos Res*. doi:[10.1016/j.atmosres.2014.11.012](https://doi.org/10.1016/j.atmosres.2014.11.012)
- Michaelides SC, Tymvios FS, Michaelidou T (2009) Spatial and temporal characteristics of the annual rainfall frequency distribution in Cyprus. *Atmos Res* 94:606–615. doi:[10.1016/j.atmosres.2009.04.008](https://doi.org/10.1016/j.atmosres.2009.04.008)
- Naoum S, Tsanis IK (2003) Temporal and spatial variation of annual rainfall on the island of Crete, Greece. *Hydrol Process* 17:1899–1922. doi:[10.1002/hyp.1217](https://doi.org/10.1002/hyp.1217)
- Shaban A (2009) Using MODIS images and TRMM data to correlate rainfall peaks and water discharges from the Lebanese coastal rivers. *J Water Resour Prot* 01:227–236. doi:[10.4236/jwarp.2009.14028](https://doi.org/10.4236/jwarp.2009.14028)
- Shrivastava R, Dash SK, Hegde MN et al (2014) Validation of the TRMM multi satellite rainfall product 3B42 and estimation of scavenging coefficients for 131I and 137Cs using TRMM 3B42 rainfall data. *J Environ Radioact* 138:132–136. doi:[10.1016/j.jenvrad.2014.08.011](https://doi.org/10.1016/j.jenvrad.2014.08.011)
- Su F, Hong Y, Lettenmaier DP (2008) Evaluation of TRMM multisatellite precipitation analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin. *J Hydrometeorol* 9:622–640. doi:[10.1175/2007JHM944.1](https://doi.org/10.1175/2007JHM944.1)
- Tan M, Ibrahim A, Duan Z et al (2015) Evaluation of six high-resolution satellite and ground-based precipitation products over Malaysia. *Remote Sens* 7:1504–1528. doi:[10.3390/rs70201504](https://doi.org/10.3390/rs70201504)
- Themistocleous K, Hadjimitsis DG, Retalis A et al (2013) Precipitation effects on the selection of suitable non-variant targets intended for atmospheric correction of satellite remotely sensed imagery. *Atmos Res* 131:73–80. doi:[10.1016/j.atmosres.2012.02.015](https://doi.org/10.1016/j.atmosres.2012.02.015)
- Tomassetti B, Verdecchia M, Giorgi F (2009) NN5: a neural network based approach for the downscaling of precipitation fields—model description and preliminary results. *J Hydrol* 367:14–26. doi:[10.1016/j.jhydrol.2008.12.017](https://doi.org/10.1016/j.jhydrol.2008.12.017)
- Tsanis I, Naoum S (2003) The effect of spatially distributed meteorological parameters on irrigation water demand assessment. *Adv Water Resour* 26:311–324. doi:[10.1016/S0309-1708\(02\)00100-8](https://doi.org/10.1016/S0309-1708(02)00100-8)
- Tsanis IK, Koutroulis AG, Daliakopoulos IN, Jacob D (2011) Severe climate-induced water shortage and extremes in Crete. *Clim Change* 106:667–677. doi:[10.1007/s10584-011-0048-2](https://doi.org/10.1007/s10584-011-0048-2)
- Ud Din S, Al-Dousari A, Ramdan A, Al Ghadban A (2008) Site-specific precipitation estimate from TRMM data using bilinear weighted interpolation technique: an example from Kuwait. *J Arid Environ* 72:1320–1328. doi:[10.1016/j.jaridenv.2007.12.013](https://doi.org/10.1016/j.jaridenv.2007.12.013)
- Verlinde J (2011) TRMM rainfall data downscaling in the Pangani Basin in Tanzania. *Master Sci Thesis Delft Univ Technol* 1:1–72
- Wang Z, Schaaf CB, Strahler AH et al (2014) Evaluation of MODIS albedo product (MCD43A) over grassland, agriculture and forest surface types during dormant and snow-covered periods. *Remote Sens Environ* 140:60–77. doi:[10.1016/j.rse.2013.08.025](https://doi.org/10.1016/j.rse.2013.08.025)
- Zargar A, Sadiq R, Naser B, Khan FI (2011) A review of drought indices. *Environ Rev* 19:333–349. doi:[10.1139/a11-013](https://doi.org/10.1139/a11-013)