PRECIPITATION ESTIMATION BASED ON SPECTRAL AND TEXTURAL FEATURES OF METEOSAT MULTISPECTRAL INFRARED DATA

Apostolos Giannakos, Haralambos Feidas

Aristotle University of Thessaloniki, School of Geology, Department of Meteorology and Climatology, 54124, Thessaloniki, Greece

Abstract

The present study aims at examining the potential of developing rainfall estimation schemes using the enhanced spectral resolution of the Meteosat Second Generation - Spinning Enhanced Visible and Infrared Imager (MSG-SEVIRI). Two different rain estimation methods were introduced that use the brightness temperature (BT) $T_{10.8}$ and brightness temperature differences (BTDs) as spectral parameters, along with textural parameters derived from the thermal infrared MSG-SEVIRI channels to estimate rainfall. The first is an algorithm based on rainfall rate assignment (RR) for each pixel of the thermal infrared MSG satellite images and the second is an Artificial Neural Network (ANN) model. Both schemes were trained using as reference spatially and temporally collocated rain data from 88 stations in Greece, for 30 rainy days covering the period from April 2008 to February 2009. Both schemes were evaluated against an independent dataset of gauge data for ten rainy days. It was found that the additional use of textural parameters along with spectral parameters improves rainfall rate estimation for both RR and ANN models in the training and validation dataset. The ANN algorithm based on both spectral and textural parameters shows the best performance among all the rainfall estimation models for both the training and validation dataset. During both the training and evaluation phase, RR model based on spectral parameters showed the worst performance among the rainfall estimation techniques. MSG satellite data and gauge rainfall measurements were used to generate precipitation estimates at various temporal scales. All rainfall retrieval algorithms overestimate the rain detected by the rain gauge network.

INTRODUCTION

Precipitation algorithms based on single satellite infrared (IR) channel data are indirectly inferred from cloud top temperatures and have limited application in the mid-latitudes, due to the difficulty of delineating rain and no rain clouds. Rainfall estimation models tend to overestimate precipitation because they are affected by the problem of rain/no rain clouds discrimination. In order to overcome the difficulty of applying a single infrared channel method in mid-latitudes, the incorporation of more than one single channel is needed. Kuhnlein et al. (2010) developed a new rainfall rate assignment method based on cloud properties derived from Meteosat multispectral satellite data in the mid-latitudes. Wei et al. (2006) displayed a multispectral convolution approach using infrared channels combined together with water vapour channel. Kurino (1997) developed a method that uses spectral cloud parameters calculated from geostationary infrared satellite data, to estimate rainfall probability and mean rainfall rate in comparison with ground radar data. The high spectral resolution of the Spinning Enhanced Visible and Infrared Imager (SEVIRI) on board the Meteosat Second Generation (MSG) satellites, with eleven 3-km-resolution channels and one 1-km-resolution visible channel provides the potential of an improved rain estimation methodology.

Several studies have shown that the performance of the rain estimation algorithm is improved when textural cloud parameters derived from MSG satellite data are used for the discrimination of precipitating cloud areas by the spatial distribution characteristics of brightness temperature levels from MSG satellite images. The spectral properties of clouds often change temporally and spatially between adjacent cloud pixels but their textural properties are often spatially distinct and tend to be less sensitive to the effects of atmospheric attenuation or detector noise. The objective of this study is to investigate the potential of using spectral and textural parameters derived from MSG-SEVIRI to improve the current IR-based rainfall retrieval algorithms that were developed from geostationary infrared data. Two different methods were proposed to develop the rainfall retrieval models. The first model is an empirical algorithm which is based on the rainfall rate assignment (RR) for each pixel of the SEVIRI multispectral infrared data and the second is a statistical approach (Artificial Neural Network, ANN) that correlates spectral and textural parameters derived from the SEVIRI dataset with rain dataset. Models' calibration and validation are carried out by comparing spectral and textural cloud parameters with rain gauge observations.

DATA AND METHODOLOGY

Data

Rain estimation schemes were trained and validated using spatially and temporally collocated 15min observation datasets of seven SEVIRI thermal infrared channels and rain gauge data for the area of Greece (Figure 1). In order to develop the two rainfall retrieval models, SEVIRI satellite data were used that include seven channels in the thermal infrared with center wavelengths at 6.2, 7.3, 8.7, 9.7, 10.8, 12.1 and 13.4 μm. They are acquired at 15 min time intervals with a spatial resolution of 3 x 3 km^2 at sub-satellite point, reaching 4 x 5 km^2 at the area of study. Rain gauge data from 88 stations of the National Observatory of Athens Hellenic (Greek) were used as reference to train the models for 30 rainy days covering the period from April 2008 to February 2009. Models were validated against ten independent rainy days that were not used for training the rainfall estimation algorithms.

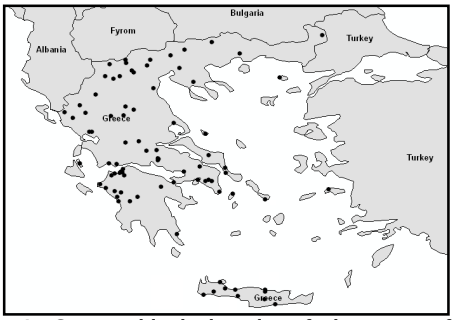


Figure 1: **Geographical domain of the area of study. Locations of rain gauge network used to train models are also shown.**

Spectral cloud parameters

The rainfall estimation techniques rely on the relationship between cloud spectral properties and physical parameters such as cloud top height, cloud optical thickness and cloud phase to detect potentially precipitating cloud areas. Clouds with large enough cloud water path together with ice particles in the upper parts have a high probability of producing precipitation (Thies et al. 2008a). The two rain retrieval schemes were developed on the basis of this relation of cloud physical parameters to spectral parameters, using the original BTDs from the SEVIRI infrared channels. The following spectral parameters were used as cloud information from the SEVIRI thermal infrared data:

Brightness temperature $T_{10.8}$, that serves as an indicator of the cloud top height. The split window $T_{10.8}$-$T_{12.1}$, being strongly related to the cloud optical thickness, can be utilized to effectively discriminate between optically thick convective clouds and optically thin cirrus clouds (Inoue 1985). Optically thick cumulus type clouds produce small BTDs due to their black-body characteristics, while optically thin cirrus clouds present large BTDs due to the differential absorption characteristics of ice crystals between the two channels.

The BTD $T_{8.7}$-$T_{10.8}$ can be utilized to gain information about the cloud phase and classify clouds as either "ice" or "water" (Thies et al. 2008a). The imaginary (absorption) component of the index of refraction,

which is a direct indicator of absorption/emission strength, differs for ice and water at these two wavelengths. More specifically, the difference in water particle absorption is small between the two wavelengths, but very large for ice particles. Radiative transfer simulations show that for ice clouds, $T_{8.7}$-$T_{10.8}$ tends to be positive in sign, whereas for low-level water clouds, $T_{8.7}$-$T_{10.8}$ tends to be small negative.

Test $T_{6.2}$-$T_{10.8}$ can effectively discriminate between high-level and mid-level clouds. Mid-level clouds produce large negative $T_{6.2}$-$T_{10.8}$ since the temperatures at the 6.2 μm channel are lower than their actual cloud top temperatures due to the absorption from water vapor above them. In contrast, $T_{6.2}$-$T_{10.8}$ usually takes very small negative values for upper level thick clouds. Positive differences may occur for deep convective clouds penetrating into the stratosphere (Fritz et al. 1993).

The BTD $T_{6.2}$-$T_{7.3}$ can be used to give information about the cloud height and about the early detection of convective activity. The two water vapor channels are dominated by atmospheric water vapor absorption. Low clouds tend to give large negative $T_{6.2}$–$T_{7.3}$ values while positive differences may occur for very high clouds reaching the stratosphere.

Test $T_{13.4}$-$T_{10.8}$ provides estimation of cloud top height and detection of cumulus cloud growth development, in MSG images. The 13.4 μm channel is the CO_2 absorption band. Low-level clouds produce large negative $T_{13.4}$-$T_{10.8}$ values since the temperatures at 13.4 μm are reduced significantly due to the absorption from CO_2 above them. $T_{13.4}$-$T_{10.8}$ difference exhibits small negative values for growing cumulus clouds since their cloud tops are above most of the CO_2 layer and ice crystal absorption is similar for both wavelengths (Mecikalski et al. 2006).

The BTD $T_{8.7}$-$T_{12.1}$ can provide information about clouds' optical thickness (Krebs et al. 2007). It takes positive values at high cloud optical thickness due to the scattering processes and the dependence on particle size which are stronger in the 8.7 μm channel compared to the 12.1 μm channel. Therefore the difference exhibits high positive values for large cloud particle size and large optical thickness. For low cloud optical thickness the BTD $T_{8.7}$-$T_{12.1}$ takes negative values because water vapor absorption in the 8.7 μm channel is higher than in 12.1 μm channel.

Test $T_{9.7}$-$T_{13.4}$ is an indicator of cloud top height. The two channels at 9.7 μm and 13.4 μm are dominated by O_3 and CO_2 absorption respectively. For high clouds over 12 km the brightness temperature at 9.7 μm show significantly larger values than brightness temperature at 13.4 μm due to the warming by stratospheric ozone (Kwon et al. 2010). Low level clouds exhibit temperatures at the 9.7 μm channel lower than their actual cloud top temperatures due to the O_3 absorption from above them.

Textural cloud parameters

The incorporation of textural parameters as complementary source of data along with spectral parameters to improve rainfall estimation models is investigated. Several recent works have shown that rainfall retrieval algorithms are significantly improved when textural measures are taken into account. Hong et al. (2004) used textural and spectral parameters calculated from GOES infrared satellite images to develop a satellite-based rainfall estimation algorithm at various temporal and spatial scales.

In this study two different statistical schemes were applied to extract textural features from MSG satellite data in the thermal infrared channel 9 at 10.8 μm: 1) Brightness Temperature Level Co-occurrence Matrix (BTLCM) and 2) Brightness Temperature Level Difference Vector (BTLDV) (Uddstrom et al. 1999, Clausi 2003). The BTLCM method is a two-dimensional histogram of brightness temperature levels for a pair of pixels which are separated by a fixed spatial relationship. It expresses the distribution density of pixels in an image. The following texture measures are calculated from the co-occurrence matrix to quantify textural variation (Haralick, 1973) and capture the spatial dependence of brightness temperature level values from the MSG satellite images: 1) Homogeneity 2) Contrast 3) Dissimilarity 4) Mean 5) Standard Deviation 6) Entropy 7) Angular Second Moment 8) Correlation. The BTLDV method is also calculated from the co-occurrence matrix but is based on absolute differences between pairs of brightness temperature levels. From the histogram of brightness temperature level differences can be computed the following measures: 1) BTLDV Angular Second Moment 2) BTLDV Entropy 3) BTLDV Mean 4) BTLDV Contrast and 5) BTLDV Inverse Difference.

DATA PREPROCESSING

The two rainfall retrieval methods were trained using rain gauge data collected for 30 rainy days from April 2008 to February 2009. Rain accumulations were recorded at 10-min time intervals and down to the rain amount of 0.1 mm. The SEVIRI data were spatially and temporally matched with rain gauge observations during the period of interest. Each 10-min rain amount recorded from gauges was first converted to rain rate and then temporally matched with the Meteosat satellite data over the corresponding 15-min scan time. The 10-min rainfall rates recorded from the rain station network were averaged over 1h, 3h, and 6h time intervals and then temporally matched with the corresponding rain estimates derived by the 15-min Meteosat scenes for each time interval. Then a geographical match of each station to a pixel location in the satellite images was performed by taking into account parallax error. The parallax can be taken into account by shifting the location of the ground station with respect to the image coordinates by a distance that depends on the satellite view angle and the cloud-top height. In this study, the cloud-top height was determined by comparing cloud-top temperatures at the 10.8 μm with the temperature profiles of the analyses of the European Center for Medium-Range Weather Forecasts (ECMWF).

For each rainy day and each station, we computed all the spectral and textural cloud parameters at the satellite temporal resolution (15 min). Furthermore the temporal aggregation of MSG satellite data was performed by applying the mean on the spectral and textural parameters values for each time scale (1h, 3h, and 6h). From the 13 texture measures that were computed from the co-occurrence matrix a calculation and comparison of the correlation coefficient was performed among all texture measures. An analysis of correlation among textures was applied to determine the best subset of textural parameters that display the least correlation. Finally four texture parameters were used in the rainfall estimation algorithms: 1) Homogeneity 2) Contrast 3) Angular Second Moment and 4) BTLDV Entropy.

Models' development methodology

Two different approaches were used to develop rainfall retrieval models. The first is an algorithm based on the rainfall rate (RR) assignment for each pixel of the satellite data and the second is an Artificial Neural Network algorithm that relates cloud textural and spectral parameters in order to estimate rain.

Rainfall Rate technique

An empirical algorithm was developed (RR1) that uses eight spectral variables $T_{10.8}$, $T_{10.8}$-T_{12}, $T_{6.2}$-$T_{7.3}$, $T_{6.2}$-$T_{10.8}$, $T_{8.7}$-$T_{10.8}$, $T_{8.7}$-$T_{12.1}$, $T_{9.7}$-$T_{13.4}$ to estimate the rainfall rate on a pixel basis. Rainfall rate algorithm (RR2) calculates rain as a function of all the twelve spectral and the four textural cloud parameters. More precisely, the rainfall rate for the two RR models is calculated as a function of the different variables x_1, x_2, …, x_i using the following equation:

$$RR(x_1, x_2 ....., x_i) = \frac{\sum rain_{obs}(x_1, x_2 ....., x_i)}{N_{rain}(x_1, x_2 ....., x_i)} \quad (1)$$

where i = 8 and 12 for RR1 and RR2 models, respectively, x_1, x_2, …, x_i denote the spectral and textural variables, N_{rain} is the number of raining pixels, and $rain_{obs}$ is the observed ground rainfall rate respectively, in each distinct interval in the (x_1, x_2, …, x_i) space. The RR parameter is based on the i-dimensional Lookup Table (LUT) that is compiled with the average rainfall rate as a function of the respective spectral and textural parameters values combinations derived from Meteosat images.

Neural Network Algorithm

A rainfall retrieval approach based on Artificial Neural Networks (ANN) was created that makes use of the high spectral resolution of the SEVIRI satellite data. The nonparametric ANN approach approximates the best nonlinear function between spectral and textural features derived from MSG satellite data and rain information from rain gauge data to estimate rainfall rate. In this study a multilayer perceptron (MLP) is selected from other statistical methods for the following reasons (Hung et al. 2009): 1) The MLP algorithm does not require any a priori knowledge of the statistical distribution of the data 2) The MLP can model non linear functions and can be trained to perform accurate generalization when inserting new unknown data 3) It is capable of performing classification with data that have high spatial variability.

First, an ANN MLP scheme (MLP1) was developed using only the eight spectral cloud parameters derived from Meteosat thermal infrared satellite images. The MLP1 optimal classification sheme consists of 8 input neurons (number of spectral cloud measures derived from SEVIRI data), 10 neurons in the hidden layer and 1 output neuron in the output layer that represents the rainfall rate estimation for each rainy pixel. Furthermore two ANN MLP algorithms (MLP1, MLP2) (Figure 2) were constructed with three layers (input, hidden, and output) that consist of 8 and 12 input neurons (number of spectral and textural cloud measures extracted from Meteosat satellite images), 10 and 15 neurons in the hidden layer and 1 output neuron in the output layer that represents the rainfall rate output for each rainy pixel.

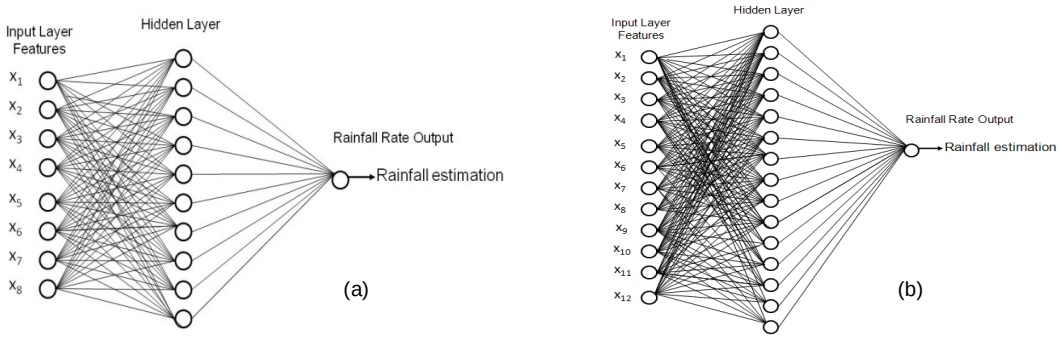


Figure 2: **Structure of Multilayer Perceptron rainfall retrieval algorithm a) MLP1 b) MLP2 that relates spectral and textural cloud parameters accordingly from SEVIRI satellite images**

RESULTS

Models' training

The following continuous statistics R (Correlation Coefficient), R^2 (Coefficient of Determination), BIAS, RMSE (Root Mean Squared Error) have been computed and used to train as well as to validate models against an independent rain gauge dataset to provide the optimal model performance in rainfall rate estimation as well as to validate models against an independent rain gauge dataset. Initially the rainfall retrieval algorithms were calibrated using spectral cloud parameters (RR1, MLP1) and spectral parameters along with textural features (RR2, MLP2) , which were calculated from a SEVIRI dataset covering a period of 30 rainy days from April 2008 to February 2009. During the training phase for each neural network algorithm, different number of neurons for the hidden layer and iterations for the adjustment of the weights were applied to find the optimal rain estimation model that minimizes the network's error. The verification scores calculated for the RR and MLP schemes are displayed in Table 1.

According to the results presented in table 1, the neural network MLP2 algorithm shows the best performance among all the four rainfall estimation models during the training phase (with a correlation coefficient ranging from to 0.26 for the 10min interval to 0.73 for the 6h interval), followed by MLP1 while the worst scores are obtained for RR1.The additional use of textural parameters along with spectral

parameters in a model improves rainfall rate estimation for both RR2, MLP2 schemes in the training phase for all different temporal scales (10min to 6h).

	15min				1h			
MODELS	**RR1**	**RR2**	**MLP1**	**MLP2**	**RR1**	**RR2**	**MLP1**	**MLP2**
R	0.08	0.12	0.21	0.26	0.20	0.25	0.34	0.38
R^2	0.01	0.02	0.05	0.07	0.04	0.06	0.12	0.15
Bias	1.30	1.28	1.24	1.26	0.53	0.77	0.95	0.80
RMSE	1.66	1.79	1.46	1.85	8.50	7.93	7.93	7.31
	3h				**6h**			
MODELS	**RR1**	**RR2**	**MLP1**	**MLP2**	**RR1**	**RR2**	**MLP1**	**MLP2**
R	0.40	0.47	0.52	0.58	0.52	0.59	0.65	0.73
R^2	0.16	0.22	0.27	0.34	0.27	0.35	0.42	0.53
Bias	0.58	0.79	1.06	0.93	0.58	0.59	0.82	0.83
RMSE	17.34	10.60	14.69	9.57	19.02	17.99	16.93	14.79

Table 1: **Verification scores computed from the algorithms during the training phase from April 2008 to February 2009**

Models' validation

The evaluation of the rainfall retrieval algorithms' performance was applied on the Meteosat satellite images selected during ten independent rainy days between April 2008 and February 2009. The results of the verification procedure are displayed in Table 2.

	15min				1h			
MODELS	**RR1**	**RR2**	**MLP1**	**MLP2**	**RR1**	**RR2**	**MLP1**	**MLP2**
R	0.05	0.09	0.19	0.22	0.17	0.23	0.27	0.34
R^2	0.002	0.01	0.04	0.05	0.03	0.05	0.08	0.11
Bias	1.18	0.87	1.25	0.90	1.18	1.45	1.27	1.67
RMSE	1.51	1.59	1.48	1.70	4.48	5.31	4.60	6.20
	3h				**6h**			
MODELS	**RR1**	**RR2**	**MLP1**	**MLP2**	**RR1**	**RR2**	**MLP1**	**MLP2**
R	0.38	0.43	0.46	0.51	0.40	0.53	0.66	0.72
R^2	0.14	0.19	0.22	0.26	0.16	0.28	0.44	0.51
Bias	0.30	0.37	1.51	1.62	0.24	0.51	0.76	0.81
RMSE	0.62	12.76	12.56	13.96	26.73	0.73	18.78	17.28

Table 2: **Verification scores computed for the four models during the evaluation phase from April 2008 to February 2009**

The comparison of the four rainfall rate estimation algorithms during the evaluation stage revealed a pattern similar to that shown during the training stage. The MLP2 algorithm still performs best among the four algorithms and the ranking of the other three algorithms according to their relative performance remains the same according to that of the dependent sample. Moreover, techniques using combined spectral and textural measures (RR2 and MLP2) provide again better results that those using spectral cloud parameters (RR1 and MLP1). Thus, MLP2 scores are best followed by MLP1 and the two RR techniques. For temporal aggregation time intervals greater than 3h a notable improvement of the rainfall estimation is apparent due to the decrease of the underestimation of very high rainfall rates. The

correspondence between observed and estimated rainfall rates increases with increasing temporal aggregation for all four algorithms.

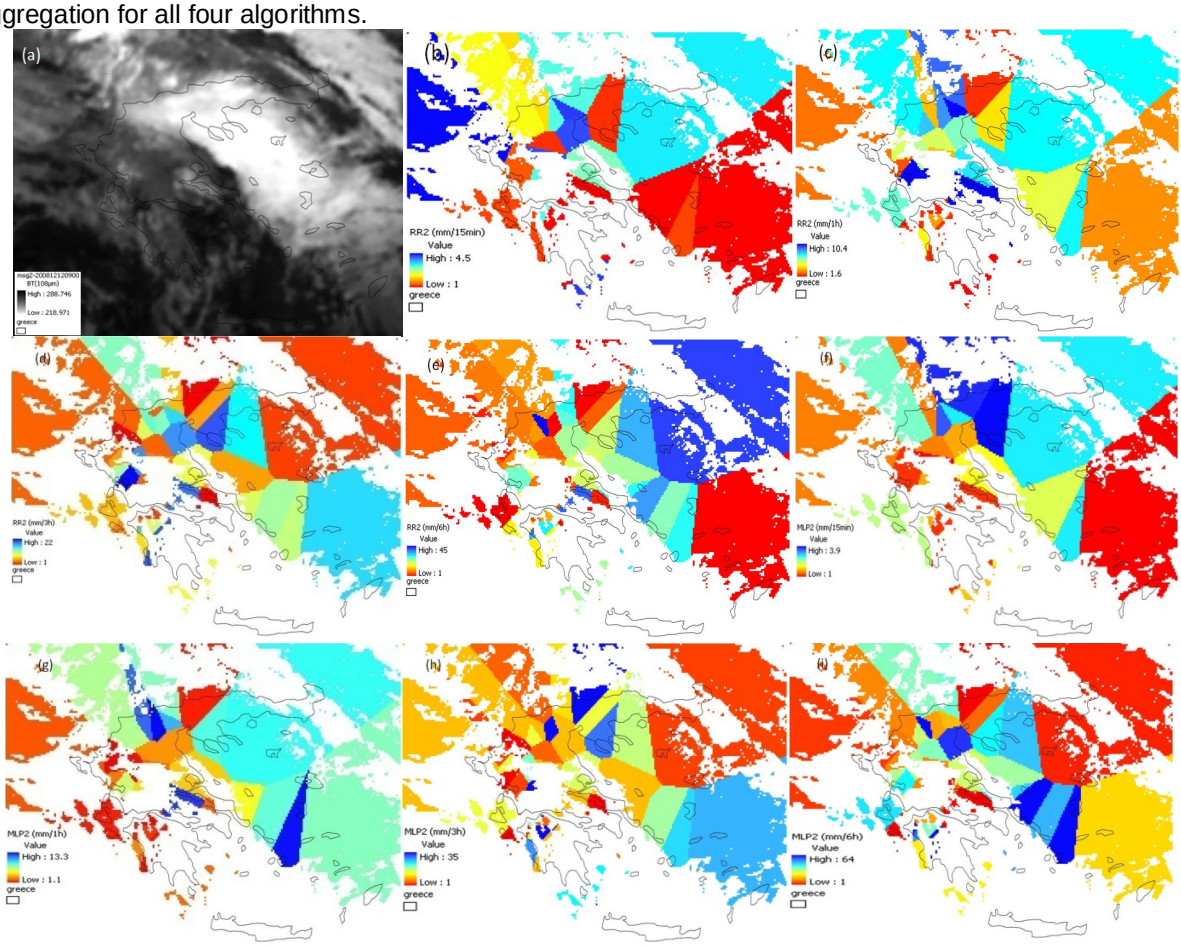


***Figure 3*: (a) SEVIRI brightness temperature at the 10.8 μm and rainfall retrieval algorithms (RR2, MLP2) (b to i) at various temporal scales (15min, 1h, 3h, 6h) for 12 December, 2008 at 09:00 UTC.**

To gain a visual impression of the spatial distribution of the rainfall rate expected amount the rain information by gauges is compared with the two rainfall estimation algorithms: 1) RR2 and 2) MLP2.

The estimated rainfall rate areas calculated by the two schemes for a scene at 09:00 UTC, 12 December, 2008, is displayed in Figure 3a to 3i for various temporal scales. Rainfall retrieval algorithms based on both spectral and textural parameters, exhibit good correspondence of estimated rainfall rate values with the cold cloud tops from the main cloud convective system. In general RR2 and MLP2 models tend to assign larger rainfall values to convective clouds but overestimate no convective precipitating cloud areas.

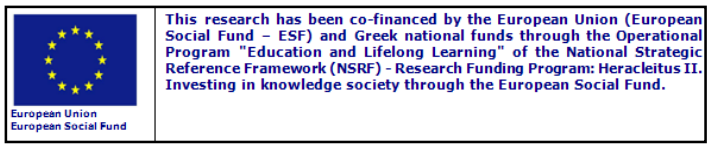
CONCLUSION

The possibility of incorporating the high infrared spectral resolution of MSG SEVIRI to estimate rainfall rate over Greece was examined. Two different schemes were developed that use spectral (RR1, MLP1) along with textural parameters (RR2, MLP2) extracted from the thermal infrared MSG channels to estimate rain. The calibration and evaluation of the rainfall retrieval algorithms based on spectral parameters showed that the developed algorithms performed poorly for short term temporal scales (15min, 1h). In contrast, encouraging results are obtained for greater temporal time intervals (3h, 6h).

However, the addition of textural cloud parameters improved rainfall rate estimation. The rainfall retrieval algorithms based on parameters on both spectral and textural parameters produced better verification scores than the algorithms based only on spectral measures. Finally, all rainfall estimation models overestimate the rainfall rate values detected by the rain gauge network.

ACKNOWLEDGEMENTS

The authors wish to thank the National Observatory of Athens for providing the rainfall rate data from their network of ground stations.



REFERENCES

Clausi D., Zhao, Y., (2003) Grey level co-occurrence integrated algorithm (GLCIA): a superior computational method to rapidly determine co-occurrence probability texture features. Comp Geosc **29**, 8337-850

Fritz, S, Laszlo, I., (1993) Detection of water vapor in the stratosphere over very high clouds in the tropics. J Geophys Res **98**, 22959-22967

Haralick, R., Shanmugan, K., Dinstein, I., (1973) Texture features for image classification. Transactions on Systems, Man, and Cybernetics, **3**, 6, 611-621

Hong, Y., Hsu, K., Sorooshian, S., Gao, X., (2004) Precipitation Estimation from Remotely Sensed Imagery Using an Artificial Neural Network Cloud Classification System. J. Appl. Meteor., **43**, 12, 1834-1853

Hung, N., Babel, M., Weesakul, S., Tripathi, N., (2009) An artificial neural network model for rainfall forecasting in Bangkok, Thailand. J. Hydrol. Earth Syst. Sci., **13**, 8, 1413-1425

Inoue, T., (1985) On the temperature and effective emissivity determination of semi-transparent cirrus clouds by bi-spectral measurements in the 10-mm window region. J. Meteor. Soc. Japan **63**, 88–99

Krebs, W., Mannstein, H, Bugliaro L., Mayer, B., (2007) A new day and night time Meteosat Second Generation Cirrus Detection Algorithm MeCIDA. J Atmospheric Chemistry and Physics **7**, 10933-10969

Kurino, T., (1997) A satellite infrared technique for estimating deep/shallow precipitation. Adv. Space Res., **19**, 3, 511–514

Mecikalski, JR., Bedka, KM., (2006) Forecasting convective initiation by monitoring the evolution of moving cumulus in daytime GOES imagery. Mon Wea Rev **134**: 49-68

Thies, B., Nauss T., Bendix J., (2008b) A new technique for detecting precipitation at mid-latitudes during daytime using Meteosat Second Generation SEVIRI. 2008 EUMETSAT Meteorological Satellite Conference, Darmstadt, Germany

Thies, B., Nauss T, Bendix J., (2008a) Discriminating raining from non-raining cloud areas at mid-latitudes using meteosat second generation SEVIRI night-time data. J. Appl. Meteor, **15**, 219-230

Uddstrom, MJ., Gray, W., (1996) Satellite Cloud Classification and Rain Rate Estimation Using Multispectral Radiances and Measures of Spatial Texture. J Appl Meteor **35**, 839-858

Wei, C., Hung, W., and Cheng, K., (2006) A multi-spectral spatial convolution approach of rainfall forecasting using weather satellite imagery. Adv. Space Res., **37**, 4, 747–753