**Advances in
Geosciences**

Statistical and neural classifiers in estimating rain rate from weather radar measurements

C. I. Christodoulou[1] and S. C. Michaelides[2]

[1]Department of Computer Science, University of Cyprus, 75 Kallipoleos Str., P.O.Box 20578, 1678 Nicosia, Cyprus
[2]Meteorological Service, 1418 Nicosia, Cyprus

Abstract. Weather radars are used to measure the electromagnetic radiation backscattered by cloud raindrops. Clouds that backscatter more electromagnetic radiation consist of larger droplets of rain and therefore they produce more rain. The idea is to estimate rain rate by using weather radar as an alternative to rain-gauges measuring rainfall on the ground. In an experiment during two days in June and August 1997 over the Italian-Swiss Alps, data from weather radar and surrounding rain-gauges were collected at the same time. The statistical KNN and the neural SOM classifiers were implemented for the classification task using the radar data as input and the rain-gauge measurements as output. The proposed system managed to identify matching pattern waveforms and the rainfall rate on the ground was estimated based on the radar reflectivities with a satisfactory error rate, outperforming the traditional Z/R relationship. It is anticipated that more data, representing a variety of possible meteorological conditions, will lead to improved results. The results in this work show that an estimation of rain rate based on weather radar measurements treated with statistical and neural classifiers is possible.

1 Introduction

Weather radars were originally (and still are) used by meteorologists in order to forecast very short-term weather conditions and issue warnings for hazardous weather phenomena. However, it was soon realized that these instruments could form potential tools in the study of a wide range of hydrological applications. Weather radars measure the electromagnetic radiation backscattered by cloud raindrops, hence their potential to estimate rainfall. Clouds that backscatter more electromagnetic radiation consist of larger droplets of rain

and therefore they can potentially produce more rain (Marshall and Palmer, 1948). The idea is to estimate rainfall rate by using weather radars instead of rain-gauges that measure rainfall on the ground. The wider spatial coverage provided by weather radars compared to that of any dense network of ground based rain-gauges is an obvious advantage. Although the idea sounds quite tempting, experience over the past four decades has revealed a series of problems related to meteorological conditions, ground clutter, shadowing by mountains, attenuation etc (Joos et al., 1998). Nevertheless, worldwide research underlines the significance of acquisition of good estimates of rain-rate with the use of weather radar (Gabella and Notarpietro, 2004; Gabella, 2004).

Traditionally, radar reflectivities are converted into instantaneous rainfall intensities by using the power-law $Z = a \times R^b$ (where Z is the radar reflectivity and R is the rain rate) (Doelling et al., 1998; Comstock et al., 2004; Rosenfeld and Amitai, 1998). A shortfall of this method is that regression-based Z/R relationships tend to overestimate the low rain intensities and underestimate the high rain intensities with the crossover at the estimated median rain volume intensity (Rosenfeld and Amitai, 1998). In this work, we explore an alternative to this methodology, namely, the use of statistical and neural classifiers for the estimation of rain rate based on weather radar reflectivity measurements (Christodoulou et al., 2004). The latter approach has recently been used to tackle a number of other meteorological and climatological problems (Michaelides et al., 2001; Christodoulou et al., 2003).

2 Material

In 1997, an experiment was carried out over the Italian-Swiss Alps during which data for two consecutive days in June and two consecutive days in August 1997 were collected. The data were simultaneously collected from the Monte
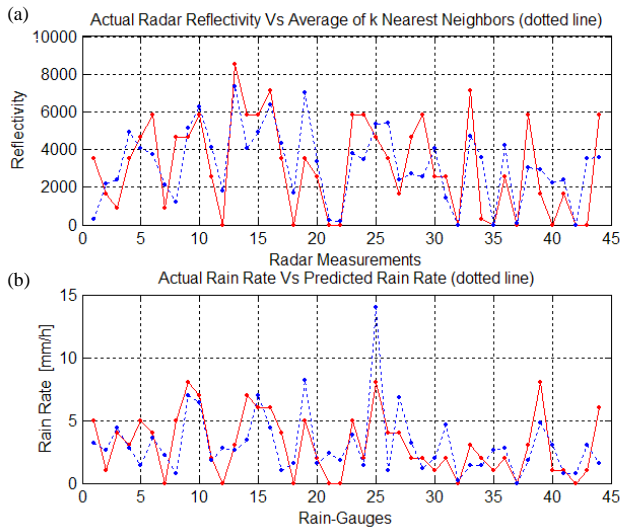
Correspondence to: C. I. Christodoulou
(cschr2@ucy.ac.cy)

(a)

(b)

Fig. 1. An example of the input evaluation radar reflectivity pattern (red continuous line) with the average of the matching radar patterns (blue dotted line) for the KNN system. Figure 1b shows the corresponding actual rain pattern (red continuous line) along with the predicted rain pattern (blue dotted line).

Lema C-band Doppler weather radar (Gabella and Notarpietro, 2004) and surrounding rain-gauges. The available data consist of radar reflectivities recorded every 5 min over 44 ground based meteorological stations and rain rates measured at these stations. For each of the above case studies, a total of 576 radar reflectivity values were available at each station (for two consecutive days 2×288=576 values). Rain-gauge measurements were taken every 10 min (i.e. only 288 values in each two day period). To make them consistent with the above 576 radar values in each case study, each 10 min rain-gauge measurement was subsequently spread over the corresponding two five-minute period, so a total of 576 rain rate values were derived, for each of the meteorological stations. The first 376 data patterns of June and the first 376 data patterns of August were used for training the system. The remaining 200 data patterns of June and 200 data patterns of August were used for evaluation. Each data pattern comprises 44 pairs of radar and corresponding rain-gauge measurements.

3 Method

For the classification task, the statistical k-nearest neighbor (KNN) (Tou and Gonzalez, 1974) and the neural network self-organizing map (SOM) (Kohonen, 1995) classifiers were used. For the implementation of the algorithms the MATLAB computing environment was used (Demuth and Beale, 1994). The KNN classification process is described in the following steps:

1. First split the available data in training and evaluation sets. In this work, 752 radar vectors were used for training and 400 vectors for evaluation as described above. For each radar vector and its 44 radar values the corresponding rain-gauge values on the ground were also known. The rain-gauge values were used as the table of truth and they were compared with the predicted by the system rain values, in order to decide whether the rain prediction using the radar measurements was successful.

2. Take the first radar vector for evaluation. Find its k (say 5) nearest neighbors from the 400 radar training vectors as the ones with the smallest Euclidean distance to the evaluation radar vector:

$$Ed_k = \sqrt{\sum_{j}^{L} \left(re_j - rt_{kj}\right)^2} \tag{1}$$

where re is the evaluation radar vector, rt_k are the training radar vectors and L (=44) is the number of the radar measurements per vector.

3. For each of the k training radar vectors selected above get the corresponding rain vector known as described in step 1. The average of the k rain vectors was the rain vector predicted by the system.

4. In order to decide whether the prediction was successful, the predicted rain vector was compared to the actual rain vector known from the rain evaluation dataset and an error rate was calculated. This was repeated for all the 400 evaluation patterns. The normalised error rate was defined as the sum of the absolute differences of the estimated to actual rain, normalised by the total actual rain as

$$Err_j = \frac{\sum_{i}^{N} |p_i - a_i|}{\sum_{i}^{N} a_i} \tag{2}$$

where p_i is the predicted rain, a_i is the actual rain and N (=400) is the number of evaluation patterns. The mean normalised error rate was calculated as

$$\overline{Err_j} = \left(\sum_{j}^{L} Err_j\right) / L \tag{3}$$

where Err_j is the error per rain-gauge j calculated above in Eq. (2), and L (=44) is the number of the rain-gauges.

Figure 1a shows an example of an evaluation radar pattern re vs. the average of the k nearest neighbor training radar

Table 1. Mean error rate of the KNN system for different values of k. The error rate was defined as the sum of the absolute differences of the predicted to actual rainfall, normalised by the total actual rainfall (Eq. 2 and 3).

k	1	2	3	4	5	10	30	50	Average
Mean Error Rate (Eq. 3)	1.041	1.058	1.041	1.049	1.027	1.010	0.974	0.985	1.023

Table 2. Mean error rate of the SOM system for different map sizes.

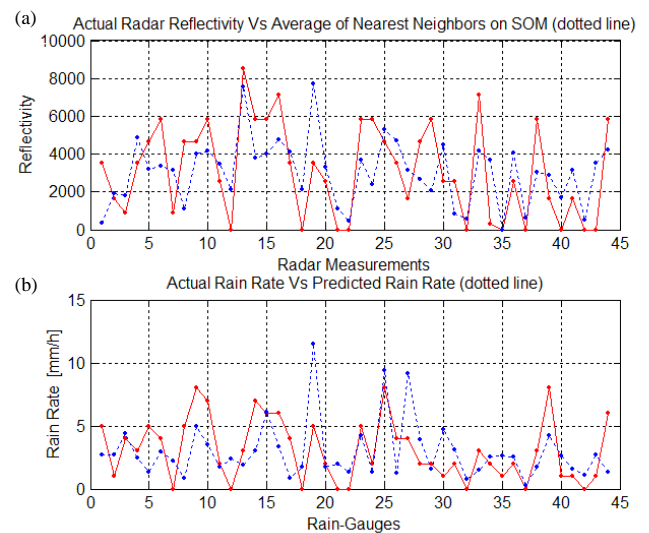| SOM	8×8	9×9	10×10	11×11	12×12	13×13	14×14	15×15	Average
Mean Error Rate (Eq. 3)	1.289	1.323	1.187	1.206	1.204	1.060	1.100	1.059	1.178

patterns rt_k. Figure 1b shows the actual rain pattern for the specific evaluation radar pattern vs. the predicted rain pattern.

In a similar way to the KNN, the neural network self-organizing map (SOM) classifier was used for classification. The SOM was chosen because it is an unsupervised learning algorithm where the input patterns are freely distributed over the output node matrix (Kohonen, 1995). The output nodes are usually ordered in a two dimensional grid, and at the end of the training phase, similar training patterns are assigned at each output node. In the evaluation phase, a test pattern is assigned to the output node with the weight vector closest to the vector of the test pattern. For the SOM rainfall prediction system, again 752 of the radar data were used for training the SOM classifier, whereas, the remaining 400 were used for evaluation. In the evaluation phase, a test radar pattern was assigned to an output node of the SOM as described above and the radar patterns, which were assigned to the specific node during training, were the radar patterns similar to the test radar pattern. For the similar radar patterns found, the corresponding rain patterns were identified from the rain data training set the same way as described for the KNN system. The average vector of the corresponding rain patterns was the predicted rainfall rate vector and the error rate was calculated as described in Eqs. (2) and (3). Figure 2 shows the same example as in Fig. 1, for the SOM system with a 13×13 output node matrix.

4 Results

Table 1 tabulates the error rate for the KNN system for different values of k and Table 2 the error rate for the SOM system for different output node matrix sizes.

The KNN classifier gave in general better results than the SOM classifier, reaching the lowest mean error rate of 0.974 when k=30. This is understandable because the problem was rather tailor-made for the KNN classifier, since it required the identification of the nearest neighbors from the pool of

Fig. 2. An example of the input evaluation radar reflectivity pattern (red continuous line) with the average of the matching radar patterns (blue dotted line) for the SOM system. Figure 2b shows the corresponding actual rain pattern (red continuous line) along with the predicted rain pattern (blue dotted line).

patterns. The average error rate was significantly improved due to a number of pattern cases where no or little rain was available, which both systems predicted correctly. The KNN results were rather consistent concerning the variations of the k value with a favour towards larger k, ranging from 1.041 error rate when k=1 to 0.985 for k=50.

For the SOM classifier best results were obtained when using the 13×13 and the 15×15 output node matrix with a mean error rate of 1.06 which was worse than the results of the KNN classifier which provided the least error rate of 0.974 when k=30. The SOM error rate decreases with larger SOM sizes and falls rather abruptly after the 12×12 SOM architecture, which shows that there is a balance when an adequate number of similar patterns are assigned to a winning
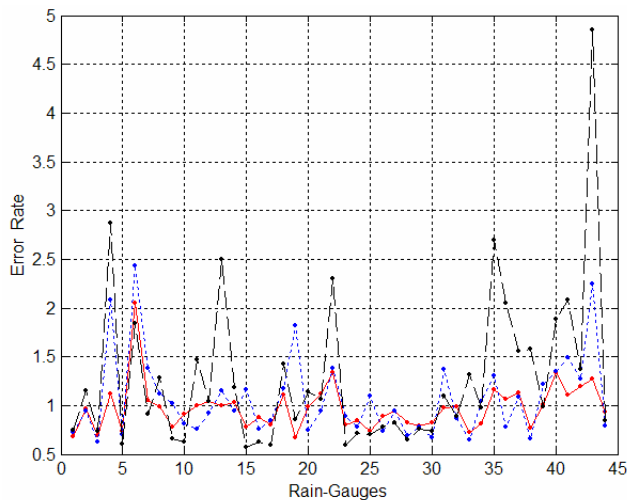
Fig. 3. Error rates calculated with Eq. (2) for the 44 rain-gauges the KNN with k=30 (red continuous line), for the SOM with a 13×13 output node matrix (blue dotted line) and the Z/R relationship (black dashed line).

node in order to provide a good matching with the rain data. The SOM classifier was trained for 10 000 epochs and the results for the SOM system are the average of three different runs in order to verify the correctness of the classification results. Test cases where the test radar pattern was assigned to an output node where no similar training radar patterns were assigned during training were ignored. Such cases are attributed to the limited number of training data, which did not cover all the possible combinations of radar reflection rain among the 44 stations. It is anticipated that more data will lead to better results.

Using the power-law $Z=a\times R^b$ for converting radar reflectivities into instantaneous rainfall intensities, with a=316 and b=1.5 (Gabella, 2004; Doelling et al., 1998; Comstock et al., 2004; Resenfeld and Amitai, 1998), yielded a mean error rate of 1.276 for the same evaluation set of 400 cases. This error rate was higher than the error rates of 1.059 obtained by the SOM system and 0.974 obtained by the KNN system. Figure 3 displays the error rates calculated with Eq. (2) for the 44 rain-gauges for the KNN, the SOM and the Z/R relationship.

5   Conclusions

A novel system was presented for the estimation of rain rate based on weather radar measurements. The system exploits the five-decade-old notion that there is a relationship between the radar measurements (reflectivities) and the rain rate (as this is measured by rain-gauges at ground level). The proposed system has the advantage that it may exploit geographical relationships and rain patterns and track repeatable patterns when they exist. The system can even tolerate miss-

ing radar measurements by finding the best matching patterns based on the available measurements and estimating rainfall rates for all the areas including the missing ones. The results of the present work suggest that the estimation of rain rate based on weather radar records and a methodology based on KNN and SOM classifiers is possible. It should be emphasised that the available data covered only two rain events covering a total of only four days. Despite that, the proposed systems managed to identify matching pattern waveforms and yielded a satisfactory error rate. It is anticipated that more data, representing a variety of possible meteorological conditions, will lead to improved results.

References

Christodoulou, C. I., Michaelides, S. C., and Pattichis C. S.: Multi-feature texture analysis for the classification of clouds in satellite imagery, IEEE Transact. Geosci. Rem. Sens., 41, 11, 2662–2668, 2003.

Christodoulou, C. I., Michaelides, S. C., Gabella, M., and Pattichis, C. S.: Prediction of Rainfall Rate Based on Weather Radar Measurements, International Joint Conference on Neural Networks, IJCNN 2004, Budapest, Hungary, 25–29 July 2004.

Comstock, K., Wood, R., Yuter, S., and Bretherton, C. : Reflectivity and rain rate in and below drizzling stratocumulus, Quart. J. Roy. Meteorol. Soc., 130, 603, Part B, 2891–2918(28), 2004.

Demuth, H. and Beale, M.: MATLAB, Neural Network Toolbox User's Guide, The MathsWork, Inc., Natick, MA. 1994.

Doelling, I. G., Joss, J., and Riedl, J. : Systematic variations of Z-R relationships from drop size distributions measured in Northern Germany during seven years, Atmos. Res., 48, 635–649, 1998.

Gabella, M. and Notarpietro, R. : Improving operational measurement of precipitation using radar in mountainous terrain – Part I: Methods, Geosci. Rem. Sens. Lett., 1, 2, 78–83, 2004.

Gabella, M.: Improving operational measurement of precipitation using radar in mountainous terrain—Part II: Verification and applications, IEEE Geosci. Rem. Sens. Lett., 1, 2, 84–89, 2004.

Joos, J., Schadler, B., Galli, G., Cavalli, R., Boscacci, M., Held, E., Della Bruna, G., Kappenberger, G., Nespor, V., and Spiess, R. : Operational use of radar for precipitation measurements in Switzerland, Final Report, VDF Hochschulverlag AG an der ETH Zurich, 1998.

Kohonen, T.: Self-Organizing Maps, Springer Series in Information Sciences, Berlin, 1995.

Marshall, J. S. and Palmer, W. M. : The distribution of raindrops with size, J. Meteorol., 5, 165–166, 1948.

Michaelides, S. C., Pattichis, C. S., and Kleovoulou, G.: Classification of rainfall variability by using artificial neural networks, Int. J. Climatol., 21, 1401–1414, 2001.

Rosenfeld, D. and Amitai, E. : Comparison of WPMM versus Regression for Evaluating Z-R Relationships, J. App. Meteorol., 37, 10, 1241–1249, 1998.

Tou., J. T. and Gonzalez, R. C.: Pattern Recognition Principles, Addison-Wesley Publishing Company, Inc., 1974.