



Investigating rainfall estimation from radar measurements using neural networks

A. Alqudah¹, V. Chandrasekar², and M. Le²

¹Yarmouk University, Irbid, Jordan

²Colorado State University, Fort Collins, Colorado, USA

Correspondence to: M. Le (leminda@engr.colostate.edu)

Received: 28 September 2012 – Published in Nat. Hazards Earth Syst. Sci. Discuss.: –

Revised: 14 December 2012 – Accepted: 10 January 2013 – Published: 4 March 2013

Abstract. Rainfall observed on the ground is dependent on the four dimensional structure of precipitation aloft. Scanning radars can observe the four dimensional structure of precipitation. Neural network is a nonparametric method to represent the nonlinear relationship between radar measurements and rainfall rate. The relationship is derived directly from a dataset consisting of radar measurements and rain gauge measurements. The performance of neural network based rainfall estimation is subject to many factors, such as the representativeness and sufficiency of the training dataset, the generalization capability of the network to new data, seasonal changes, and regional changes. Improving the performance of the neural network for real time applications is of great interest. The goal of this paper is to investigate the performance of rainfall estimation based on Radial Basis Function (RBF) neural networks using radar reflectivity as input and rain gauge as the target. Data from Melbourne, Florida NEXRAD (Next Generation Weather Radar) ground radar (KMLB) over different years along with rain gauge measurements are used to conduct various investigations related to this problem. A direct gauge comparison study is done to demonstrate the improvement brought in by the neural networks and to show the feasibility of this system. The principal components analysis (PCA) technique is also used to reduce the dimensionality of the training dataset. Reducing the dimensionality of the input training data will reduce the training time as well as reduce the network complexity which will also avoid over fitting.

1 Introduction

Rainfall on the ground is dependent on the four dimensional structure of precipitation aloft. Scanning radar observations can capture the four dimensional structure of precipitation. However, it is difficult to express the relation between the radar observations and ground rainfall in a simple form. The key challenge in radar rainfall estimation is the space-time variability in precipitation microphysics, such as drop size distribution (DSD) and its impact on rainfall on the ground. It is well established that an empirical Z – R relation is not sufficient to capture the variability and has large uncertainty and it needs to be adaptively adjusted based on validation (Cifelli and Chandrasekar, 2010). Prior research has shown that neural networks can be used to estimate ground rainfall from radar measurements (Xiao and Chandrasekar, 1997; Xiao et al., 1998; Liu et al., 2001; Orlandini and Morlini, 2000). The usefulness of the rainfall estimation using neural networks is subject to many factors such as the representativeness and sufficiency of the training dataset, the generalization capability of the network to new data, seasonal changes, regional changes, and so on. An artificial neural network (ANN), often simply called a neural network (NN), is a nonparametric method to establish the nonlinear mapping from input space to a target space. It consists of interconnected group of neurons, each characterizing a simple function.

Neural Network techniques have been used in weather radar applications such as rainfall and snowfall estimation. In addition they have been used for rain profile classification. Neural network based radar snowfall estimation was introduced first by Xiao and Chandrasekar (1996). Rainfall estimation was introduced by the same authors (Xiao and

Chandrasekar, 1995). An attempt to do rain type classification using self-organizing maps (SOM) was introduced by Zafar et al. (2003).

Radial Basis neural network is capable of learning a complex functional relation from high dimension input space to the target space. It has been demonstrated in prior work that RBF Neural Network is capable of learning the relation between ground radar measurements and rain gauge data (Liu et al., 2001; Orlandini and Morlini, 2000; Xu and Chandrasekar, 2005; Teschl et al., 2007). In this paper, an adaptive relation between ground radar measurements and rain gauge measurements will be developed in the training process, and studies are conducted to improve the performance of the network. One of the major challenges in building estimators using neural network is to choose the appropriate input. While it is clear that the rainfall estimate depends on the full 3-D structure of precipitation aloft, using the full 3-D data as input creates a demand for enormous training process. The principal components analysis technique is used to modify the input to rainfall estimation neural network. Data from Melbourne, Florida NEXRAD ground radar (KMLB) and a network of gauges from the years 2006, 2007, 2008 and 2009 are used to demonstrate the neural network based radar rainfall estimation. The performance of radar rainfall estimation will be analyzed and compared against rain gauge measurements. The improvement due to PCA filtering is quantitatively analyzed. This paper is organized as follows: Sect. 2 introduces the radial basis function neural network for radar rainfall estimation, whereas Sect. 3 describes the corresponding adaptive network. In Sect. 4 the various options of the vertical profiles are explored. The input structure to the neural network is discussed in Sect. 5, while Sect. 6 summarizes the important results.

2 Radial basis function (RBF) neural network for rainfall estimation

The radial basis function (RBF) network is part of the multilayer feed forward neural network (MLF-NN) class. It gets its name from the use of the radial basis function as activation function in the hidden layer. Figure 1 shows the structure of an RBF network (Liu et al., 2001). It contains three layers which are the input layer, the hidden layer and the output layer. The input vectors are fed to the input layer where they pass to the hidden layer. The hidden layer units or neurons have nonlinear radial-basis functions where each has its own center vector and width or size. The output of each neuron is calculated based on the Euclidean distance between the input vector and the center vector of that neuron. The outputs of the hidden layers are weighted and added linearly at the output layer.

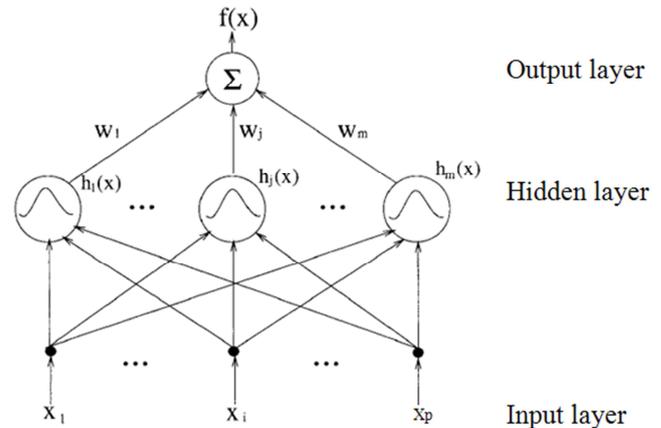


Fig. 1. The general structure of RBF network.

2.1 RBF neural network architecture

As mentioned above, the RBF NN has three layers (input, hidden, and output layer). The input layer accepts the input vector $\mathbf{X} = [x_1, x_2, \dots, x_p]^T$. The hidden layer consists of m neurons with $h(\mathbf{x})$ as transfer function. In this work, $h(\mathbf{x})$ was chosen to be the Gaussian RBF given by

$$h_j(\mathbf{x}) = \exp \left[- \sum_{i=1}^p \frac{(x_i - c_{ij})^2}{r_{ij}^2} \right] \quad (1)$$

and the output $f(\mathbf{x})$ can be calculated by a linear combination of the hidden layer outputs as follows:

$$f(\mathbf{x}) = \sum_{j=1}^m w_j h_j(\mathbf{x}), \quad (2)$$

where $\mathbf{c}_j = [c_{1j}, c_{2j}, \dots, c_{pj}]^T$ is the center vector of neuron j , $\mathbf{r}_j = [r_{1j}, r_{2j}, \dots, r_{pj}]^T$ is the size or width vector of neuron j , m is the number of neurons in the hidden layer, and w_j is the weight from neuron j to the output layer.

2.2 Input/target of the RBF neural network

NEXRAD radars also known as WSR-88D (the Weather Surveillance Radar-1988 Doppler) radar operate at S-band frequency with beam at around 0.88–0.96 degree. The antenna of the NEXRAD radar scans continuously in a sequence of 360° in the azimuth at various elevation angles. The antenna has two different scan patterns and they are done alternatively based on the status of the atmosphere (Precipitation/No Precipitation). KMLB radar is one of the WSR-88D radars that located at Melbourne, FL NEXRAD site (28.113° N, 80.654° W). Radar data and rain gauge observations that are used to train and test this RBFNN were collected during the years 2006, 2007, 2008 and 2009 over Melbourne, Florida area. Averaged precipitation from year 2006

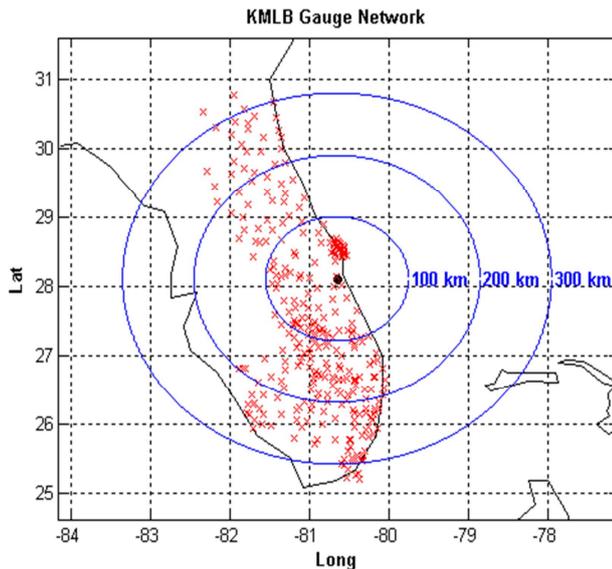


Fig. 2. KMLB gauge network.

to 2009 shows similar trend over Melbourne, Florida site. During summer time from June to September, rainfall rate was the largest in the year. The total precipitation was gradually increasing from 42 to 55 inches in the period 2006 to 2009. Radar data (radar reflectivity factor) will be used as an input to the neural network and the rain gauge corresponding to that input will be the target of the neural network. Radar data were obtained from the radar Constant Altitude Plan Position Indicator (CAPPI) datasets. The PPI data were collected in a volume and transformation technique was used to map the data on a Cartesian grid. Subsequently, constant altitude datasets were selected for the analysis. The lowest height level of the CAPPI scans is 1 km and the highest level is 4 km. The spacing between the CAPPI levels is chosen to be 1 km. The gauge data were maintained by NASA TRMM program. Around KMLB radar, the gauge networks that were considered are Kennedy Space Center (KSC), South Florida Water Management District (SFL), and St. Johns Water Management District (STJ). Within a 100-km radius around KMLB site, these networks have 33, 46 and 99 rain gauges, respectively, which accumulated rain every 5 min. Figure 2 shows a geographical map of radar and rain gauges used in this study. The radar measurement of interest in this work is only radar reflectivity factor Z_h at the horizontal polarization. CAPPI data containing Z_h values at 1 km, 2 km, 3 km, and 4 km in height with 1-km horizontal resolution as shown in Fig. 3.

2.3 Adaptive RBF neural network for ground radar rainfall estimation

The target in this paper is to estimate rainfall every volume scan or every six minutes. At the end of a day, and if we

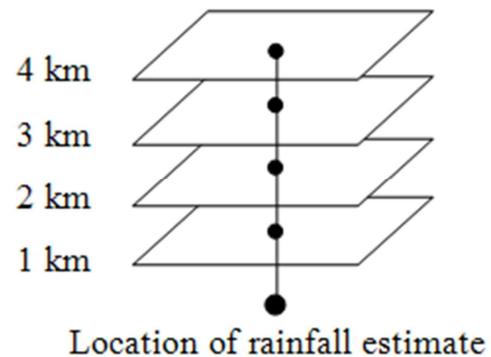


Fig. 3. Data locations for 4 point input vector.

have new data available, we need to create a new model to efficiently train the neural network. A neural network is designed to estimate rainfall from the ground radars measurements as input and the corresponding rain gauge measurements as a target. This neural network is trained adaptively and weights are updated on daily basis. The vertical reflectivity profiles are taken starting at 1 km and going up to 4 km with 1-km vertical resolution. The rain gauge measurements were averaged over six minutes and considered as a target for the network.

Since our goal is to estimate rainfall every six minutes, over a year, the network might get very large and hard to train from the beginning if we keep adding neurons every time we have new input. Another concern is that the new data might not carry new information. Therefore, the idea of adaptively training the RBFNN on a daily basis is useful (Liu et al., 2001). To include the information from the new data, it is necessary to update the network not only by adding some neurons, but also by removing some neurons. If the new data carries similar input data with different output, there is no need to retrain the network again; rather we just need to recalculate the weights from the hidden units to the output unit. This process reduces the complexity of the network, and the redundancy of the data, and by doing this it improves the generalization of the network, and reduces the training time because adjusting the weights is a simple operation and that would make the operation faster (Liu et al., 2001). Figure 4 shows the concept of daily adapting the neural network.

3 Performance of the adaptive rbf neural network

3.1 Training the neural network

The radar data used in this evaluation was collected at the Melbourne, Florida site. The neural network was trained adaptively at the end of every day. The target of the network was the rain gauge measurements that were collected from the tipping bucket rain gauge networks around those three radars. The data were from years 2006, 2007, 2008 and

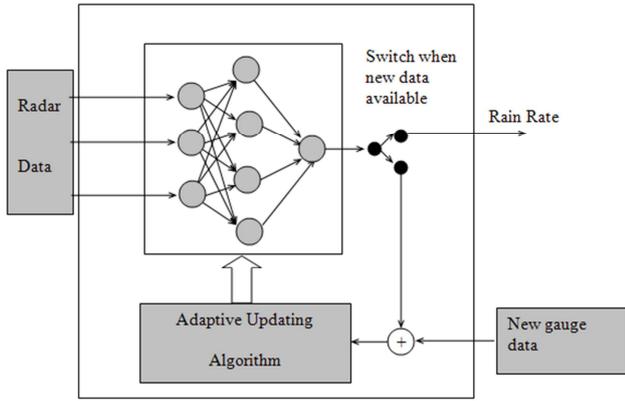


Fig. 4. Dynamic Neural network.

2009. Data were taken within 100 km around the radar. Input training data which was the radar measurements was taken at 1 km 2 km, 3 km and 4 km in height as shown in Fig. 3. This would make the size of the input vector to be four ($p = 4$). Rain gauge data were averaged over six minutes to meet the radar sweep time. Figure 5 shows a representation of how the neural network is trained.

3.2 Testing and validating the neural network

At the end of any day, once the network is trained, and then for the following day when we have new observations available, this data is used to estimate rain rate using the neural network that was trained. The estimation was validated against the rain gauge measurements of that day. The dashed line of Fig. 5 shows a schematic diagram of the rain rate estimation using the trained neural network.

3.3 Performance evaluation

The performance of the network was calculated using the following metrics:

$$\text{FracBias} = \frac{\frac{1}{N_g} \sum_{i=1}^{N_g} (\text{RFn}(i) - \text{RFg}(i))}{\overline{\text{RFg}}} \tag{3}$$

$$\text{Corr} = \frac{\frac{1}{N_g} \sum_{i=1}^{N_g} (\text{RFg}(i) - \overline{\text{RFg}})(\text{RFn}(i) - \overline{\text{RFn}})}{\sigma(\text{RFg})\sigma(\text{RFn})} \tag{4}$$

$$\text{NSE} = \frac{\frac{1}{N_g} \sum_{i=1}^{N_g} |\text{RFg}(i) - \text{RFn}(i)|}{\overline{\text{RFg}}} \tag{5}$$

$$\text{NRMSE} = \frac{\left(\frac{1}{N_g} \sum_{i=1}^{N_g} (\text{RFg}(i) - \text{RFn}(i))^2 \right)^{1/2}}{\overline{\text{RFg}}} \tag{6}$$

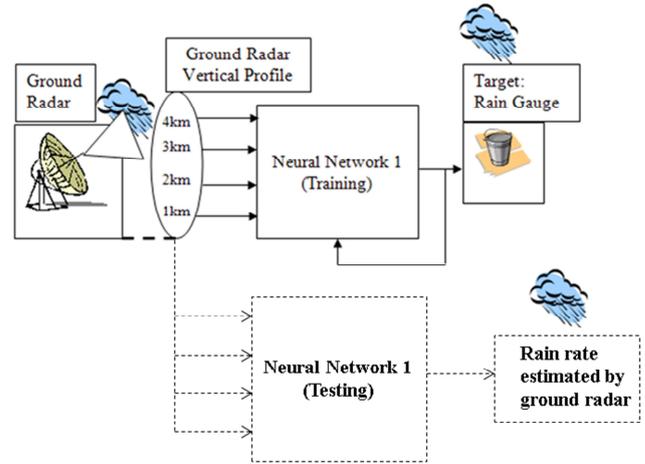


Fig. 5. Neural Network training. Input ground radar reflectivity. Target: Rain gauge. Rain rate estimation (dashed line). Input ground radar reflectivity. Output: Rain rate estimation.

where FracBias, Corr, NSE, and NRMSE are fractional bias, correlation, normalized standard error, and normalized root mean square error, respectively. RFn and RFg denote the estimated rainfall and the actual rain gauge, respectively, and Ng is the size of the data. The network performance was also compared with the simple Z–R relation used in NEXRAD radars, and with the best-fit against gauge.

Table 1 shows hourly rainfall accumulation scores of the adaptive neural network using data from 2006 to 2009 over KMLB. As it can be seen in the table, the performance of the neural network approach is much better than the performance of the Z–R relation ($Z = 300 R^{1.4}$). It is also shown that the performance of the neural technique is very close or better than the performance of the best-fit method even though the fitting was done “after the fact”. The best-fit method was based on finding the coefficients (a, b) of $Z = aR^b$ that would best-fit Z and R. The fitting was done based on least square approximations. As we see, the Z–R relation has significant bias compared to the rain gauge measurements, while the neural network product very small bias.

Table 1 also shows that the correlation and the NRMSE scores of the neural networks are better than that for the Z–R relation. The neural networks score higher correlation and lower NRMSE, while the Z–R scores lower correlation and higher NRMSE which means a lower variation from truth (rain gauge) in the favor of the neural network technique. The proposed technique has good scores compared to the best-fit method as well. As we see, the neural network scores are either very close or sometimes better than the best-fit scores taking into consideration again that the best-fit was done after the fact. Figures 6, 7, 8, 9 and 10 show the same conclusions that can be inferred from the table. The figures show better scatter and standard deviation plots of the neural network performance when compared to the best-fit plots. The

Table 1. Performance evaluation of the NN rain rate estimation, the Z–R estimation, and the best-fit estimation against rain gauge. Data from year 2006 to 2009 over KMLB. (Hourly rainfall accumulation).

| Algorithm | FracBias (%) | | | | NRMSE (%) | | | |
|------------------------------|--------------|-------|-------|-------|-----------|------|------|------|
| | 2006 | 2007 | 2008 | 2009 | 2006 | 2007 | 2008 | 2009 |
| NN Est. vs. Rain Gauge | 2.4 | 3.1 | 2.7 | 2.4 | 44.2 | 45.9 | 40.9 | 38.9 |
| Z–R Est. vs. Rain Gauge | –44.9 | –44.0 | –44.5 | –39.6 | 68.8 | 68.4 | 74.4 | 58.3 |
| Best-fit Est. vs. Rain Gauge | –26.5 | –16.6 | –19.9 | –17.3 | 51.5 | 46.8 | 45.8 | 40.1 |
| | Corr. | | | | NSE | | | |
| NN Est. vs. Rain Gauge | 0.81 | 0.82 | 0.84 | 0.84 | 0.30 | 0.31 | 0.27 | 0.27 |
| Z–R Est. vs. Rain Gauge | 0.72 | 0.75 | 0.65 | 0.80 | 0.50 | 0.51 | 0.52 | 0.45 |
| Best-fit Est. vs. Rain Gauge | 0.80 | 0.82 | 0.83 | 0.85 | 0.36 | 0.32 | 0.30 | 0.29 |

Table 2. Annual correlation between rain gauge and radar reflectivity at different heights (1 to 10 km).

| Height (km) | KMLB | KMLB | KMLB | KMLB |
|-------------|------|------|------|------|
| | 2006 | 2007 | 2008 | 2009 |
| 1 | 0.62 | 0.57 | 0.50 | 0.52 |
| 2 | 0.63 | 0.56 | 0.49 | 0.51 |
| 3 | 0.60 | 0.56 | 0.45 | 0.46 |
| 4 | 0.55 | 0.52 | 0.39 | 0.38 |
| 5 | 0.52 | 0.49 | 0.37 | 0.39 |
| 6 | 0.55 | 0.50 | 0.42 | 0.44 |
| 7 | 0.50 | 0.45 | 0.40 | 0.42 |
| 8 | 0.45 | 0.42 | 0.35 | 0.40 |
| 9 | 0.42 | 0.38 | 0.34 | 0.39 |
| 10 | 0.39 | 0.36 | 0.32 | 0.37 |

figures also show a comparative performance of the neural network approach when compared to the Z–R approach.

4 Effect of radar measurement height profiles on rain rate estimation using neural networks

In the previous results, the neural networks were designed and tested based on radar measurements taken up to 4 km in height starting at 1 km with 1-km spacing. In this section, we investigate the effect of the height going from 4 km up to 9 km keeping the same spacing. In other words, we need to find the answer to whether radar measurements for heights lower or higher than 4 km would improve the performance of the network or not?

To answer this question, we first calculated the correlation between the rain gauge measurements and the radar reflectivity factor measured at different heights starting at 1 km and up to 10 km. It can be seen from Table 2 that the correlation is higher for heights less than or equal to 4 km for most of the years. This result is not surprising considering the average melting level in Melbourne area is between 4 to 5 km.

Table 3. The effect of using radar measurements from different heights on the performance of the NN rain rate estimator. Data from year 2006 to 2009 over KMLB. (Hourly rainfall accumulation).

| Height (km) | FracBias (%) | | | | NRMSE (%) | | | |
|-------------|--------------|------|------|------|-----------|------|------|------|
| | 2006 | 2007 | 2008 | 2009 | 2006 | 2007 | 2008 | 2009 |
| 4 | 2.4 | 3.1 | 2.7 | 2.4 | 44.2 | 45.9 | 40.9 | 38.9 |
| 5 | 1.6 | 3.1 | 4.3 | 4.2 | 44.8 | 46.1 | 44.2 | 41.6 |
| 6 | 2.7 | 1.7 | 3.9 | 5.2 | 46.7 | 46.4 | 50.3 | 45.6 |
| 7 | 3.4 | 1.9 | 4.6 | –8.1 | 50.6 | 50.2 | 51.8 | 66.2 |
| 8 | 9.8 | 5.9 | 10.7 | 11.2 | 58.9 | 54.3 | 56.9 | 56.3 |
| 9 | 15.1 | 5.5 | 14.9 | 20.3 | 68.2 | 60.9 | 62.2 | 71.5 |
| | Corr. | | | | NSE | | | |
| 4 | 0.81 | 0.82 | 0.84 | 0.84 | 0.30 | 0.31 | 0.27 | 0.27 |
| 5 | 0.80 | 0.81 | 0.82 | 0.82 | 0.30 | 0.31 | 0.29 | 0.29 |
| 6 | 0.80 | 0.80 | 0.77 | 0.79 | 0.31 | 0.32 | 0.31 | 0.30 |
| 7 | 0.77 | 0.77 | 0.76 | 0.54 | 0.33 | 0.33 | 0.33 | 0.42 |
| 8 | 0.73 | 0.74 | 0.76 | 0.74 | 0.40 | 0.35 | 0.37 | 0.38 |
| 9 | 0.68 | 0.69 | 0.75 | 0.69 | 0.45 | 0.39 | 0.41 | 0.47 |

In continuation of answering the previous question, the neural networks were trained and tested using rain gauges and radar measurements up to different heights (4 km to 9 km). Table 3 shows the results of this test over KMLB site. The neural networks were trained and tested starting at 1 km and going up to the height shown in the first column in the tables, with km vertical spacing. The table shows that when the radar measurements were taken from 1 km and up to 4 km in height, the performance was better than that if we take radar measurements up to heights higher than 4 km for most of the cases. This result was also observed by Li et al. (2003). It was found that equispaced input from 1 km to 4 km in height above the gauge would give the best result.

The reason why the correlation was higher for heights up to 4 km is that the rain region was within 4 km in height as shown in Fig. 11 of radar data from year 2009, and going higher to the melting layer and to the ice region will make the correlation between the gauge on the ground and the melting layer and the ice region (which are represented by the radar

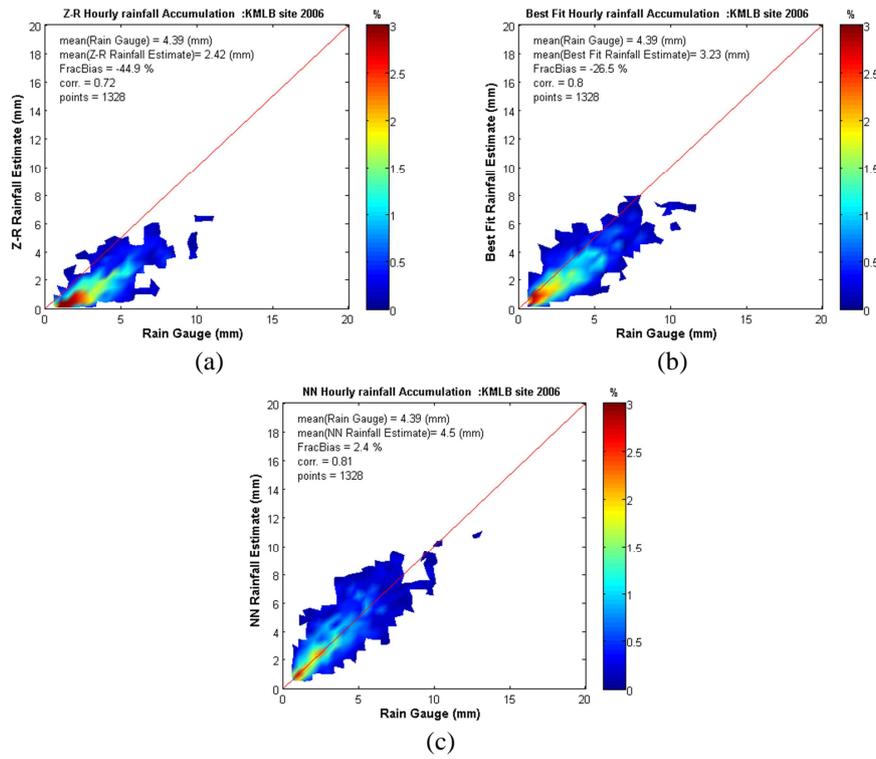


Fig. 6. Actual rain gauge vs. (a) Z–R estimate, (b) best-fit estimate and (c) NN estimate. Data from year 2006 over KMLB. (Hourly rainfall accumulation).

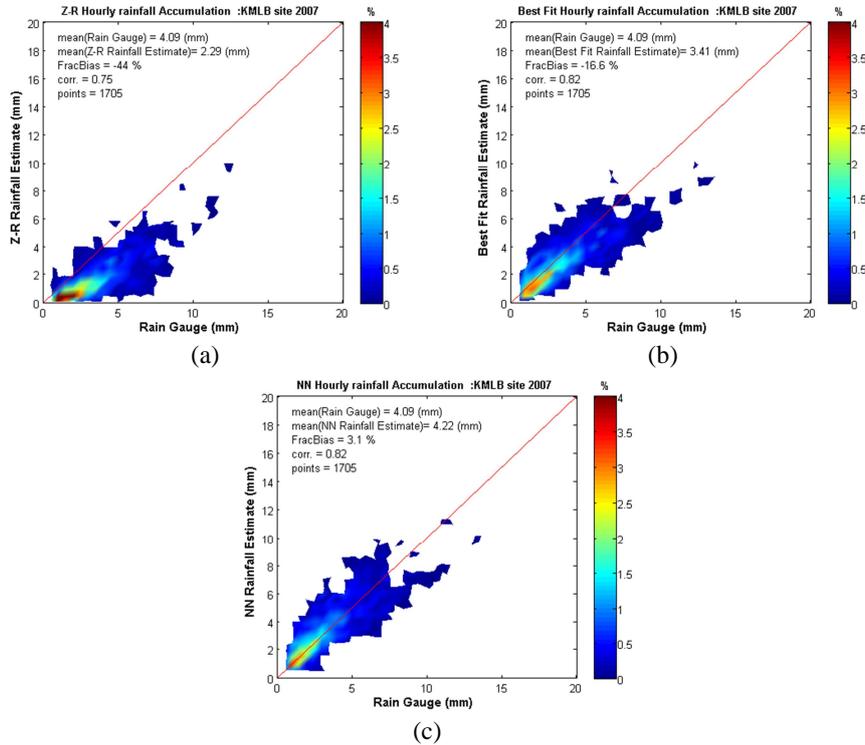


Fig. 7. Actual rain gauge vs. (a) Z–R estimate (b) best-fit estimate and (c) NN estimate. Data from year 2007 over KMLB. (Hourly rainfall accumulation).

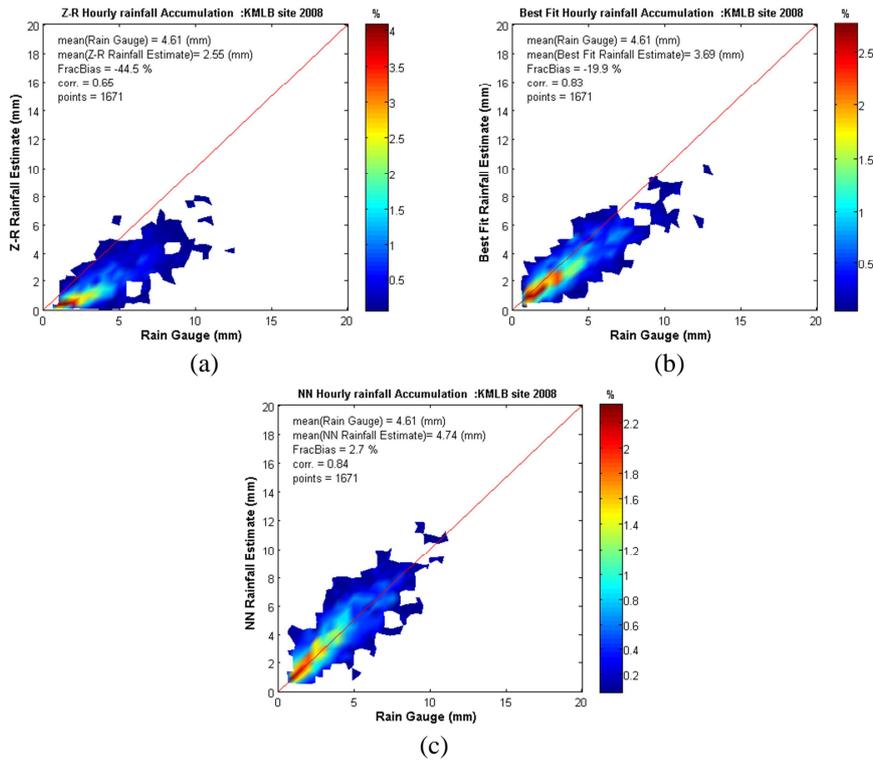


Fig. 8. Actual rain gauge vs. (a) Z–R estimate (b) best-fit estimate and (c) NN estimate. Data from year 2008 over KMLB. (Hourly rainfall accumulation).

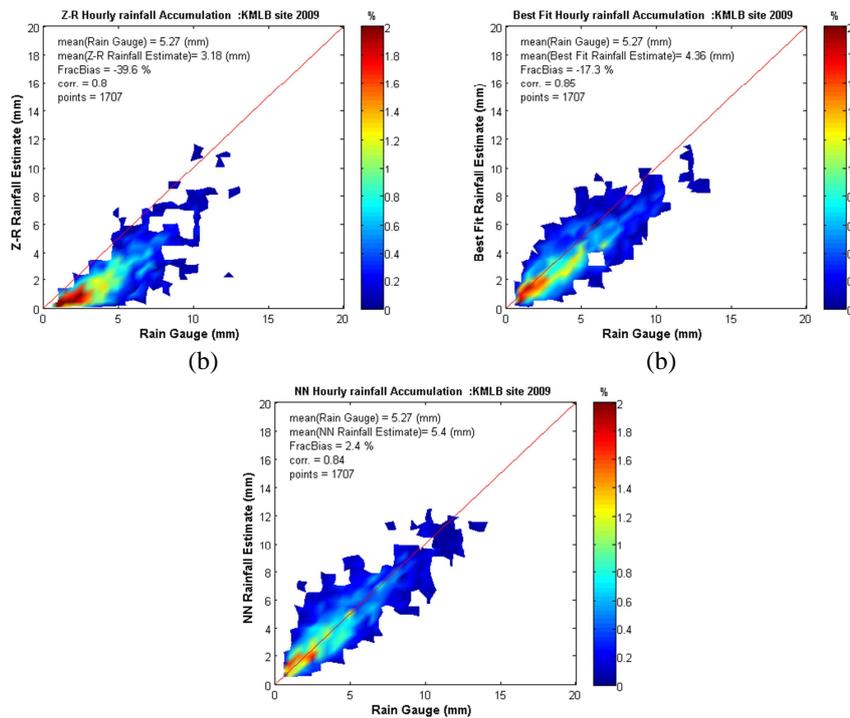


Fig. 9. Actual rain gauge vs. (a) Z–R estimate (b) best-fit estimate and (c) NN estimate. Data from year 2009 over KMLB. (Hourly rainfall accumulation).

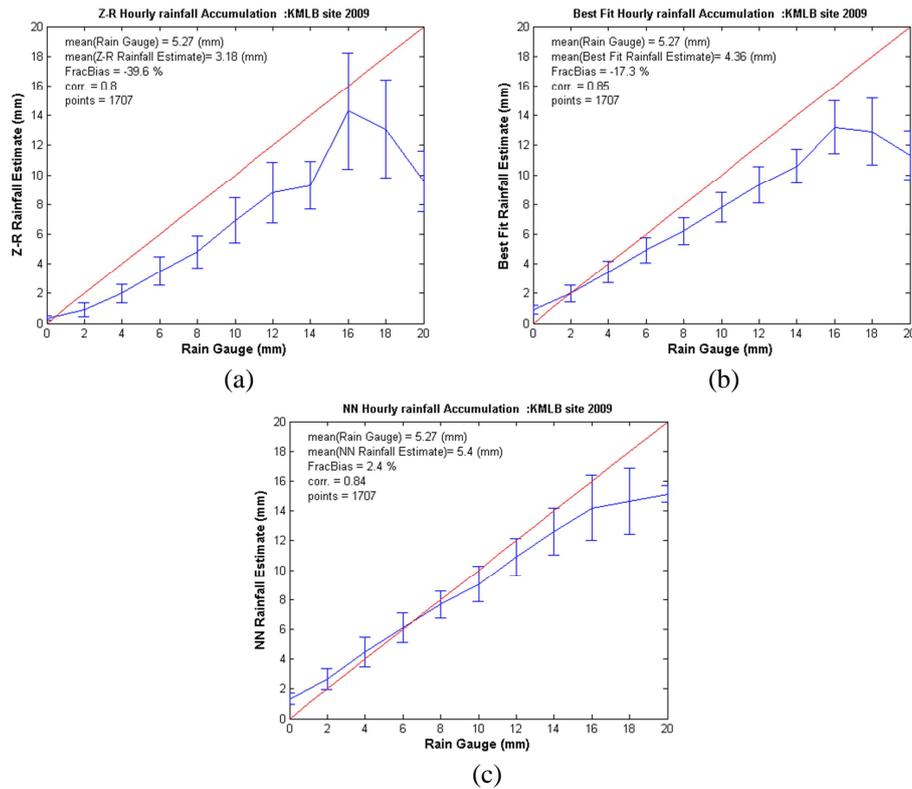


Fig. 10. Standard deviation plot of actual rain gauge vs. (a) Z–R estimate (b) best-fit estimate and (c) NN estimate. Data from year 2009 over KMLB. (Hourly rainfall accumulation).

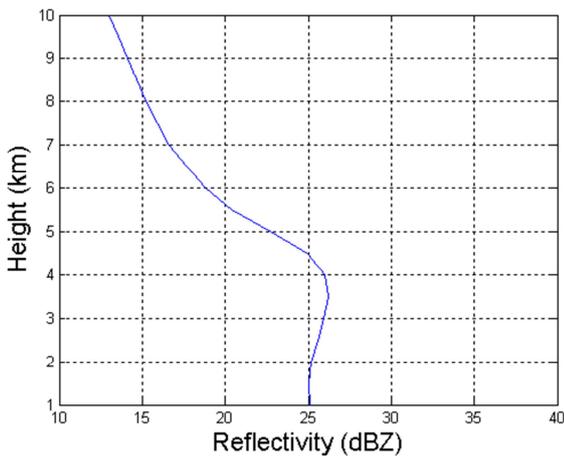


Fig. 11. Radar Reflectivity Factor vs. Height. Data from year 2009.

reflectivity factor) to be smaller than that for lower heights. Therefore, it will be easier for the neural network to find the relation between the rain gauges and the radar reflectivity factor if measurements at height up to 4 km are used.

It is worth mentioning that taking radar measurements higher than 4 km will reduce the number of good (valid)

profiles that can be used to train the network; this is because of low rain rate measurements are mostly related to weak storms, which usually do not have measured reflectivity at higher altitudes. Therefore, considering measurements at higher altitudes would eliminate weak storms from being included in the analysis.

5 Using principal component analysis to reduce the input size to the rainfall NN

5.1 Principal component analysis

In this section, the input radar reflectivity factor Z along 4 CAPPI levels is explored by applying the principal component analysis (PCA) over the standardized values of Z_i ($i = 1, \dots, 4$), where Z_i represents the radar reflectivity factor measured at height i . The standardized values of Z_i (\tilde{Z}_i) are given by

$$\tilde{Z}_i = (Z_i - E[Z_i]) / \sqrt{\text{Var}[Z_i]}, \quad (7)$$

where $E[Z_i]$ and $\text{Var}[Z_i]$ denote the sample mean and variance of Z_i . Standardization is necessary due to the different range values Z_i might have. If we define $S\tilde{z}$ to be the sample covariance matrix of \tilde{Z} , with elements $(S\tilde{z})_{ij}$ given by

Table 4. The performance of the RBF NN using radar data versus 2 PCs as input. (Hourly rainfall accumulation).

| Site/Year | FracBias (%) | | Corr. | | NSE | | NRMSE (%) | | Network Size (Centers) | |
|-----------|--------------|-------|------------|-------|------------|-------|------------|-------|------------------------|-------|
| | Radar data | 2 PCs | Radar data | 2 PCs | Radar data | 2 PCs | Radar data | 2 PCs | Radar data | 2 PCs |
| KMLB 2006 | 2.4 | 0.9 | 0.81 | 0.82 | 0.30 | 0.29 | 44.2 | 43.7 | 34 | 15 |
| KMLB 2007 | 3.1 | 2.6 | 0.82 | 0.80 | 0.31 | 0.33 | 45.9 | 47.9 | 62 | 19 |
| KMLB 2008 | 2.7 | 1.9 | 0.84 | 0.83 | 0.27 | 0.28 | 40.9 | 42.1 | 37 | 15 |
| KMLB 2009 | 2.4 | 3.4 | 0.84 | 0.84 | 0.27 | 0.27 | 38.9 | 38.1 | 22 | 18 |

$$\begin{aligned}
 (S\tilde{z})_{ij} &= \text{Cov}[\tilde{Z}_i, \tilde{Z}_j] \\
 &= \sum_{s=1}^M (\tilde{Z}_i)_s (\tilde{Z}_j)_s / M : (i = 1, \dots, 4; j = 1, \dots, 4) \quad (8)
 \end{aligned}$$

where M is the number of input patterns, then we need to find the eigenvectors e and the eigenvalues D_{ii} of the covariance matrix as seen below:

$$e^{-1} S\tilde{z} e = D_{ii}. \quad (9)$$

After that we calculate the principal components (PCs) that are associate with the eigenvalues using

$$PC_1 = e_1^T \tilde{Z}, PC_2 = e_2^T \tilde{Z}, \dots, PC_4 = e_4^T \tilde{Z}. \quad (10)$$

The goal of using the PCA concept in this context is to reduce the dimensionality of the training data to a level where we still can get good performance. In this section, we train the neural network using the principal components rather than the radar reflectivity factor Z . To get benefit from this concept and to reduce the dimensionality of the training data, we are going to neglect those principal components with small eigenvalues. There are two methods to decide which principal components to neglect. The first one is to sum the eigenvalues from the largest to the lowest, and when the sum exceeds a certain threshold we stop adding eigenvalues, and we use only those whose eigenvalues were considered in the addition. Another way to find out which principal components to include is to use the Fisher’s Maximum Coverage Test (Mielke and Berry, 2007).

5.2 Performance evaluation of RBF NN using PCA

The PCA technique was applied to the data from years 2006, 2007, 2008 and 2009 over the KMLB site. It was found that two principal components were enough to provide reasonable performance than that using four levels of radar reflectivity factor values to train the network. Two input configurations were tested in this regard, in each one the performance of the neural network was measured as well as the time it takes the neural network to train, and the number of neurons needed (network size). The purpose of including the training

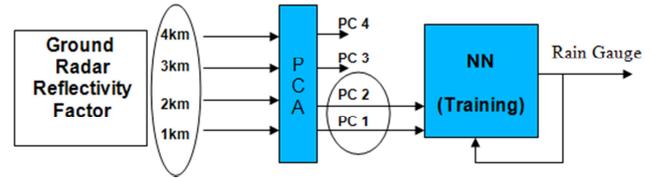


Fig. 12. Schematic of NN trained with input configuration 2.

time and the network size is to see the effect of the size of the input data and to see how feasible the network can be in order to be applied in real time.

5.2.1 Input configuration 1

This input configuration is the same one used in the previous evaluations. The network was trained using radar measurements at 1, 2, 3 and 4 km in height, and rain gauges were the target. The purpose of including this configuration is to estimate the training time and the network size at each case in order to find out the improvement brought by PCA technique. Figure 4 shows the configuration of the input where radar reflectivity factor at four different heights was used to train the network. The performance of the neural network using this input configuration is shown in Table 4 and it will be compared to the performance of the next input configuration.

5.2.2 Input configuration 2

In this configuration, the network was trained using the PCs calculated from the radar measurements at 1, 2, 3 and 4 km in height. Only two principal components were used in the training together with their corresponding rain gauges. The chosen PCs were those whose eigenvalues accumulation is more than the threshold value chosen. Figure 12 shows a schematic of this configuration and Table 4 also shows the neural networks performance when using this input configuration. As we see in Table 4, the performance of the neural networks based rainfall estimate is improved. Although training time is dependent on the computer, the major reduction was seen in the training time with almost 50 % less than the time spent using the previous input configuration. In addition, the other performance metrics, such as FracBias,

Corr., NSE, and NRMSE, were almost the same in most of the cases. Another improvement of using this configuration is the reduction of the network complexity; the previous network was designed using 4-D input vectors, while this network is designed using 2-D input vectors. As can be seen from the table, the network size got reduced by about 50% in most of the cases. This reduction is very important especially when the network is going to be implemented in real time.

6 Summary

The main goal of this paper is to investigate neural network based rainfall estimation from ground radar. Radial basis function neural network was the main neural network architecture applied to do the estimation. The main approach was based on a neural network that is designed based on rain gauges and vertical profile of ground radar measurements. The ground radar was used in this regard is KMLB NEXRAD radar. Three rain gauge networks around this radar, namely KSC, SFL, STJ, were used in comparison. The following points summarize the results of this paper:

- A neural network technique was used to estimate rainfall from ground radar measurements. The effect of the radar vertical profile height on rainfall estimation was examined. It was found that measurements up to 4 km were giving better performance in most of the cases, for the Melbourne region
- The neural network performance was compared with the $Z-R$ relation and with a statistical approach (best-fit) against the rain gauge. It was found that the neural network performance was better in most of the cases. The $Z-R$ relation was underestimating the rain rate and was unable to capture the storm variations in most of the cases.
- The “principal component analysis” (PCA) was used to reduce the input size. Two principal components were used to train the neural network. Significant improvements were achieved in computation time while maintaining the statistical metrics (FracBias, correlation, NSE, and NRMSE) in comparison to full training.

Acknowledgements. This research was supported by the NASA PMM program.

Edited by: G. Boni

Reviewed by: two anonymous referees

References

- Cifelli, R. and Chandrasekar, V.: Dual polarization radar rainfall estimation in rainfall: State of the Science, edited by: Testik, F. Y. and Gebremichael, M., Am. Geophys Union, Washington, 105–125, 2010.
- Li, W., Chandrasekar, V., and Gang, G. Xu.: Investigations in Radar Rainfall Estimation Using Neural Networks, *Int. Geosci. Remote Sens. Sym.*, 4, 2347–2349, 2003.
- Liu, H., Chandrasekar, V., and Gang, G. Xu.: An Adaptive Neural Network Scheme for Radar Rainfall Estimation from WSR-88D Observations, *J. Appl. Meteorol.*, 40, 2038–2050, 2001.
- Mielke, P. and Berry, K.: *Permutation Methods, A distance Function Approach*, New York, Springer, 2nd Edn., 439 pp., 2007.
- Orlandini, S. and Morlini, I.: Artificial Neural Network Estimation Of Rainfall Intensity From Radar Observations, *J. Geophys. Res.*, 105, 24849–24861, 2000.
- Teschl, R., Randeu, W. L., and Teschl, F.: Improving weather radar estimates of rainfall using feed-forward neural networks, *Comput. Int. Earth Environ. Sci.*, 20, 519–527, 2007.
- Xiao, R. and Chandrasekar, V.: Multiparameter Radar Rainfall Estimation Using Neural Network Techniques, *Conference on Radar Meteorology*, 27th, Vail, CO, USA, 9–13 Oct. 1995, 199–201, 1995.
- Xiao, R. and Chandrasekar, V.: Multiparameter Radar Snowfall Estimation Using Neural Network Techniques, *Proceedings IGARSS 1996*, 566–568, 1996.
- Xiao, R. and Chandrasekar, V.: Development of a Neural Network Based Algorithm for Rainfall Estimation from Radar Observations, *IEEE Trans. Geosci. Remote Sens.*, 35, 160–171, 1997.
- Xiao, R., Chandrasekar, V., Liu, H., and Gorgucci, E.: Detection of Rain/No Rain Condition on ground From Radar Data Using A Kohonen Neural Network, *Geosci. Remote Sens. Symp. Proc.*, 1998, 159–161, 1998.
- Xu, G. and Chandrasekar, V.: Operational Feasibility of Neural-Network-Based Radar Rainfall Estimation, *IEEE Geosci. Remote Sens. Lett.*, 2, 13–17, 2005.
- Zafar, B., Mubarak, K., and Chandrasekar, V.: Evaluation of Precipitation Type Determination from TRMM Observation, *Proc. IEEE Conf. Geoscience and Remote Sensing*, 2003, 3175–3177, 2003.