A Machine Learning Model for Radar Rainfall Estimation based on Gauge Observations

Haiming Tan, V. Chandrasekar, and Haonan Chen
Department of Electrical and Computer Engineering
Colorado State University
Fort Collins, Colorado, USA
Haiming.Tan@colostate.edu

***Abstract*—Rainfall estimation based on radar measurements has been addressed by using parametric algorithms such as Z-R relation. However, such empirical relations may not be sufficient to capture the space-time variability of precipitation. In this paper, we introduce a DMLP-based machine learning model for rainfall estimation by using multiple layers to capture the complex abstractions of radar reflectivity at different attitude levels. The radar data collected by the Weather Surveillance Radar – 1988 Doppler (WSR-88DP) in Melbourne, Florida (i.e., KMLB radar) are used for demonstration purposes, while the rain gauge data are used for training purposes. The rainfall product derived from the DMLP model is compared against an independent rain gauge dataset, which shows excellent performance of the new machine learning based rainfall model.**

I. INTRODUCTION

Rainfall estimation based on radar measurements has been pursued for several decades. Generally, this research problem is addressed by using parametric algorithms such as reflectivity-rainfall rate relations (i.e., Z-R relations). In recent years, the dual-polarization based rainfall algorithms have been fairly successfully (e.g., Cifelli et al. 2011; Chen and Chandrasekar 2015; Chen et al. 2016). However, the empirical relations may not be sufficient to capture the space-time variability of precipitation microphysics in term of Drop Size Distribution (DSD). On the other hand, Machine Learning, a nonparametric method, can be used to estimate rainfall based on radar observations. This method, which takes into account of both radar observations and rainfall measurements from ground rain gauges, has been demonstrated in a number of previous studies (e.g., Xiao and Chandrasekar 1997; Liu et al. 2001) for rainfall rate estimation. However, the traditional machine learning-based rainfall estimation is limited in practice due to the simple model structure. This paper approaches the same problem using a new machine learning method based on Deep Multiplayer Perceptron (DMLP) model. DMLP is widely used by the community of image processing and speech recognition. Compared to traditional neural network based machine learning approach, the DMLP model has larger number of hidden layers and more complex architecture for data representation. The high level structured information and knowledge can be extracted from the low level data features through the hierarchical learning process.

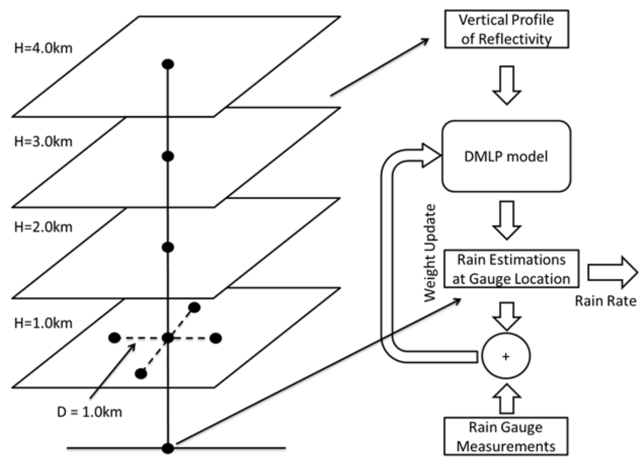

Figure 1. Conceptual diagram of the DMLP-based machine learning model for radar rainfall estimation.

In this paper, we introduce a DMLP-based machine learning model for rainfall estimation using radar reflectivity measurements. This model is designed to capture the complex abstractions of reflectivity at different attitude levels by using multiple layers feature identification and extraction. The high level abstraction can be used for rainfall estimation independently or fused with other precipitation data products. For demonstration purposes, the radar data collected by the Weather Surveillance Radar – 1988 Doppler (WSR-88DP) in Melbourne, Florida (i.e., KMLB radar) are used. In addition, the rain gauge data are used for training purposes. Figure 1 shows the conceptual diagram of the proposed machine-learning model. The performance of rainfall product derived based on the new machine learning model is then compared against independent rain gauge observations. This paper will also explore the dimensionality and convergence for various input models.

II. PRELIMINARY RESULT

In this section, KMLB radar data collected during the storm events in 2005 were selected for implementation of the DMPL model. This radar dataset includes a large number of convective and stratiform rainfall cases. Radar data in polar coordinates are first mapped to Constant Altitude Plan Position

Indicator (CAPPI) at multiple vertical levels with a spatial resolution of 1km by 1km. A four-layer DMPL model is designed to estimate rainfall rates with the CAPPI data from 1km to 4km vertical levels (see also Figure 1). In this study, rainfall measurements from three rain gauge networks are used for training and validating purposes, including Kennedy Space Center (KSC), South Florida Water Management District (SFL), and St. Johns Water Management District (STJ), which include 33, 46, and 99 rain gauge stations, respectively. It should be mentioned that in order to keep the independence between training and testing, the rain gauge data are randomly divided into two datasets. One is used for training the DMPL model, while the other is used for validating the radar rainfall products.

Figure 2 shows a scatter plot of radar rainfall estimates based on the DMPL model versus the rainfall observations from the validation gauge stations. For peer comparison, a traditional Z-R relation, which is commonly applied to the National Weather Service (NWS) radars, is implemented from:

$$Z = 300R^{1.4} \tag{1}$$

where Z is radar reflectivity in linear scale, R is rainfall rate in mm/hr. Figure 3 illustrates the scatter plot of rainfall estimates using equation (1) versus rainfall observations from the validation gauges. Comparing the preliminary results in Figures 2 and 3, it is concluded that the proposed DMPL model has great potential for rainfall estimation. In order to further evaluate the performance of the DMPL model, the bias (*Bias*) and root-mean-square-error (*RMSE*) are computed, which are respectively defined as follows:

$$Bias = < R_G - R_R > \tag{2}$$

$$RMSE = \sqrt{\langle (R_G - R_R)^2 \rangle} \tag{3}$$

where R_R and R_G denote rainfall observations from KMLB radar and validation gauge, respectively; the angle brackets stand for sample average.

The biases of radar rainfall estimates based on the DMPL model and Z-R relation are 0.16 mm/hr and 3.89 mm/hr, respectively, whereas the *RMSE*s are 5.55 mm/hr and 15.31mm/hr, respectively. Obviously, the DMLP model has much better performance than Z-R relation in terms of at *Bias* and *RMSE*.

Acknowledgement

This research was primarily supported by the NASA GPM/PMM program. The authors also acknowledge the support from the National Science Foundation (NSF) Accelerating Innovation Research (AIR) program and the Hazards SEES program. The Florida rain gauge data used in this research are managed by Kennedy Space Center (KSC), South Florida Water Management District (SFL), and St. Johns Water Management District (STJ).

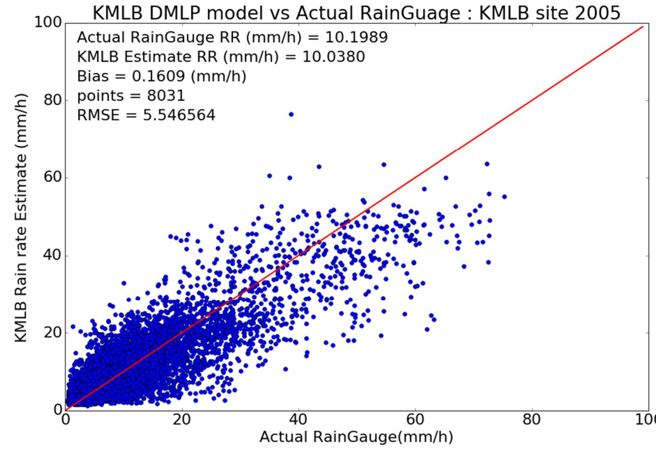

Figure 2. Radar rainfall estimates based on the DMLP model vs. rainfall measurements from validation gauges. The data includes a large number of rainfall events in 2005.

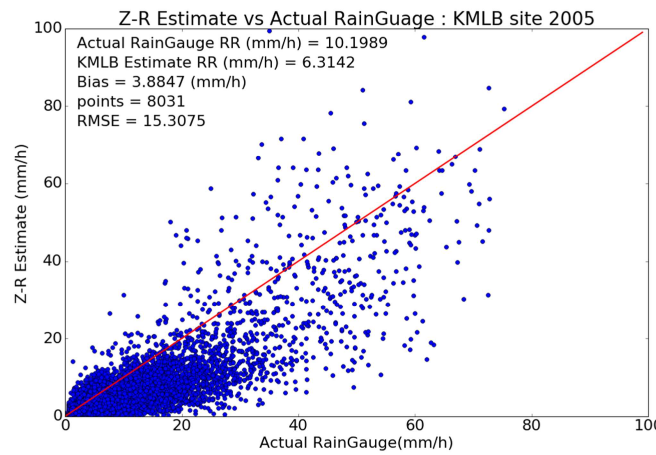

Figure 3. Radar rainfall estimates based on Z-R relation in equation (1) vs. rainfall measurements from validation gauges. This dataset includes the same rainfall events with those shown in Figure 2.

References

[1] Xiao, R. and V. Chandrasekar. 1997: Development of a neural network based algorithm for rainfall estimation from radar measurements. *IEEE Trans. Geosci. Remote Sens.*, **36**, 160-171.

[2] Liu, Hongping, V. Chandrasekar, and G. Xu, 2001: An Adaptive Neural Network Scheme for Radar Rainfall Estimation from WSR-88D Observations. *J. Appl. Meteor.*, **40**, 2038–2050.

[3] Cifelli, R., V. Chandrasekar, S. Lim, P. C. Kennedy, Y. Wang, and S. A. Rutledge, 2011: A new dual-polarization radar rainfall algorithm: Application in Colorado precipitation events. *J. Atmos. Oceanic Technol.*, **28**, 352–364.

[4] Chen, H., and V. Chandrasekar, 2015: Estimation of Light Rainfall Using Ku-Band Dual-Polarization Radar. *IEEE Trans. Geoscience and Remote Sensing*, **53(9)**, 5197-5208.

[5] Chen, H, V. Chandrasekar, and R. Bechini, 2016: An Improved Dual-Polarization Radar Rainfall Algorithm (DROPS2.0): Application in NASA IFloodS Field Campaign. *Journal of Hydrometeorology*. (in press).