Water Resources Research

TECHNICAL REPORTS: METHODS
10.1029/2019WR026255

# Deep Learning for an Improved Prediction of Rainfall Retrievals From Commercial Microwave Links

**Jayaram Pudashine[1], Adrien Guyot[1], Francois Petitjean[1], Valentijn R. N. Pauwels[1], Remko Uijlenhoet[2], Alan Seed[3], Mahesh Prakash[4], and Jeffrey P. Walker[1]**

[1]Department of Civil Engineering, Monash University, Melbourne, Victoria, Australia, [2]Hydrology and Quantitative Water Management Group, Wageningen University and Research, Wageningen, The Netherlands, [3]Bureau of Meteorology, Victoria, Australia, [4]Data 61, CSIRO, Melbourne, Victoria, Australia

**Key Points:**
- A novel approach is proposed to estimate rainfall from only maximum and minimum attenuation data from microwave links using a deep learning
- The RNN model trained and tested using disdrometer data outperformed existing rainfall estimation methods from microwave link attenuation
- This disdrometer-trained model also outperformed rainfall estimation methods when applied to Commercial Microwave Link data

**Correspondence to:**
J. Pudashine,
jayaram.pudashine@monash.edu

**Citation:**
Pudashine, J., Guyot, A., Petitjean, F., Pauwels, V. R. N., Uijlenhoet, R., Seed, A., et al. (2020). Deep learning for an improved prediction of rainfall retrievals from commercial microwave links. *Water Resources Research*, *55*, e2019WR026255. https://doi.org/10.1029/2019WR026255

Received 1 SEP 2019
Accepted 25 MAY 2020
Accepted article online 31 MAY 2020

**Abstract** Commercial microwave links (CMLs) have proven useful for providing rainfall information close to the ground surface. However, large uncertainties are associated with these retrievals, partly due to challenges in the type of data collection and processing. In particular, the most common case is when only minimum and maximum received signal levels (RSLs) over a given time interval (hereafter 15 min) are stored by mobile network operators. The average attenuation and the corresponding rainfall rate are then calculated based on a weighted average method using the minimum and maximum attenuation. In this study, an alternative to using a constant weighted average method is explored, based on a machine learning model trained to produce actual attenuation from minimum/maximum values. A rainfall retrieval deep learning model was designed based on a long short-term memory (LSTM) model architecture and trained with disdrometer data in a form that is comparable to the data provided by mobile network operators. A first evaluation used only disdrometer data to mimic both attenuation from a CML and corresponding rainfall rates. For the test data set, the relative bias was reduced from 5.99% to 2.84% and the coefficient of determination ($R^2$) increased from 0.86 to 0.97. The second evaluation used this disdrometer-trained LSTM to retrieve rainfall rates from an actual CML located nearby the disdrometer. A significant improvement in the overall rainfall estimation compared to existing microwave link attenuation models was observed. The relative bias reduced from 7.39% to 1.14% and the $R^2$ improved from 0.71 to 0.82.

**Plain Language Summary** A deep learning model was designed and trained using a disdrometer-derived data set and further applied to retrieve rainfall from commercial microwave link data. This model showed significant improvements in rainfall estimation over a constant weighted average method.

## 1. Introduction

It is now widely recognized that a wealth of environmental data can be acquired from sources that were never intended for such a purpose (Gazit & Messer, 2018). The use of commercial microwave links (CML) as precipitation measurement sensors is one example. This technology has proven useful for providing rainfall data, with several examples around the world, including the Netherlands (Leijnse et al., 2007; Overeem et al., 2016b), Israel (Goldshtein et al., 2009; Messer et al., 2006), Germany (Chwala et al., 2012; Graf et al., 2019; Smiatek et al., 2017), Switzerland (Bianchi et al., 2013), Czech Republic (Fencl et al., 2013), Africa (Gosset et al., 2016), Brazil (Rios Gaona et al., 2018) and Pakistan (Sohail Afzal et al., 2018). Rainfall estimation from such CML networks is based on attenuation of the transmitted electromagnetic signal as it passes through rain.

The density of CML networks is expected to expand globally (Uijlenhoet et al., 2018). The major advantage of such link-derived rainfall data above other conventional sources of rainfall information is that it provides a path-integrated measurement over hundreds of meters to several kilometers close to the ground. Consequently, this can provide a complementary source of rainfall information, especially for countries where there are limited ground observational networks. Such rainfall information would be particularly important for urban areas where there is a need for high temporal and spatial resolution of the data to represent the fast dynamics of the hydrological system (Kavetski et al., 2011) and where it is not easy to install gauges due to buildings and vandals. There have been several examples of using this rainfall

©2020. American Geophysical Union. All Rights Reserved.

Printed by [Monash University - 049.17.091.252 - /doi/epdf/10.1029/2019WR026255] at [24/09/2020].

measurement for hydrological modeling (Brauer et al., 2016; Smiatek et al., 2017), flood early warning systems (Eshel et al., 2017), and even the improvement of rainfall estimates from existing rainfall measurement sensors (Liberman et al., 2014).

The CML-derived rainfall retrieval accuracy and temporal resolution are dependent on a number of factors, one of which is the way the received signal level (RSL) is stored by the telecommunication companies or operators (Leijnse et al., 2008). Operators use a network management system (NMS) to collect and store RSL data in their cellular network for quality control monitoring of their network. In most cases, the minimum and maximum values are stored for a 15-min window (Messer, 2018; Uijlenhoet et al., 2018). Usually, this 15-min time window and parameters (minimum, maximum RSL, and transmit power) stored by the NMS are hard-coded by the hardware provider, which is usually sufficient for network quality monitoring purposes. However, parameters for the rainfall retrieval using the power relationship, derived from (typically 30-s) drop size distribution (DSD) data and applied to 15-min minimum and maximum or average attenuation could lead to uncertainty. Also, within this 15-min interval, the exact time when the minimum and maximum RSL reading occurred is unknown (Ostrometzky & Messer, 2014), which could lead to uncertainties in the rainfall retrieval. Several studies tried to make use of such available data sets to retrieve rainfall (Ostrometzky & Messer, 2014, 2017; Overeem et al., 2011, 2013, 2016a). Using only minimum and maximum RSL values, Overeem et al. (2011) employed an extra coefficient to provide the estimate of time-averaged rain rate over the 15-min sampling period. This coefficient provides the relative contribution of minimum and maximum attenuation ($\alpha$ and 1 $\alpha$ for minimum and maximum, respectively) to obtain the weighted average attenuation. They suggested $\alpha$ as 0.33 based on a set of data for the Netherlands with the same value used for all time steps. Thus, this $\alpha$ remains uncertain as the distribution of the rainfall or attenuation is usually not consistent across any 15-min period. Similarly, Ostrometzky and Messer (2017) also used various extreme value distribution functions for estimating time-average rain rates from minimum and maximum signal levels.

Theoretically, there is no physical limit of a NMS on recording interval, meaning that it is possible to record the average RSL, poll the data at higher frequency or even record in real time, except for the limitation of storage capacity or data transfer (Chwala & Kunstmann, 2019; Messer, 2018; Uijlenhoet et al., 2018). However, as these cellular networks have been designed and optimized for providing efficient telecommunication rather than measuring rainfall, the opportunistic use for rainfall monitoring needs to deal with the existing scheme of data sampling and storage. Thus, there is a need for a robust and more accurate methodology of estimating accurate rainfall using only the minimum and maximum RSLs that are currently recorded.

Recent developments in the field of deep learning, especially in the branch of deep neural networks offer a great opportunity to model various physical processes from data, especially if large quantities of data are available (Reichstein et al., 2019). As rainfall occurrence and RSL are time-dependent phenomena, and as a prediction is required at each time step, recurrent neural networks (RNNs) are well-suited: They can learn time dependency and are applicable to series of varying lengths (Shi et al., 2014). Such networks have been successfully used in other fields of computer science, such as speech recognition, language translation, and video and motion prediction (Graves & Schmidhuber, 2009; Mathieu et al., 2015). Closer to the application presented here, Mishra et al. (2018) have implemented a deep learning framework to distinguish dry and wet periods from communication satellite data to improve rainfall retrievals. Recently, Habi and Messer (2018) and Polz et al. (2019) have also used machine learning techniques for wet-dry classification of commercial microwave links. Similarly, there have been a few other studies using deep learning for rainfall runoff modeling (Hu et al., 2018; Kratzert et al., 2018), but according to an extensive search this is the first work that employs a recurrent neural network for improving the rainfall estimation from commercial microwave link data.

The primary objective of this study was to design and apply a deep learning model for improving rainfall estimation using CML data. The specific objectives of this study were (1) to train and validate a deep learning model using a disdrometer data set and (2) to use this model to predict rainfall using limited data (only minimum and maximum RSL data) from a CML. To achieve these objectives, laser-based disdrometer data were collected for more than 1 year and then used for training, validation, and testing. Subsequently, the disdrometer-trained model was used to retrieve rainfall from a CML situated in the proximity of the disdrometer. A range of deep learning model architectures were evaluated and compared with the existing

Printed by [Monash University - 149.171.091.252 - /doi/epdf/10.1029/2019WR026255] at [24/09/2020].

approach (in this case the result obtained using weighted average method). In the overall rainfall retrieval process, all other steps are based on the steps followed by Overeem et al. (2016a), including baseline estimation, wet antenna attenuation, and attenuation-rainfall relationship.

## 2. Data

### 2.1. Disdrometer Data

The disdrometer data used in this study were obtained from an OTT PARSIVEL[1] disdrometer installed at Mount View Reservoir; Glen Waverley [37°53′24″S, 145°10′23″E] located 20 km from Melbourne CBD (Central Business District) in Southeast Australia. This instrument measures the drop size and fall speed of the hydrometeors distributed into 32 non-equidistant classes (Jaffrain & Berne, 2011; Jaffrain et al., 2011). From the raw drop size diameter and velocity data, rainfall intensity, reflectivity, and other parameters such as visibility can be derived according as outlined in Loffler-Mang and Joss (2000). Disdrometer data covering the period from December 2017 until November 2018 were used for this study, with a measurement time interval of 30 s. The raw data collected from the OTT Parsivel[1] were processed following the same steps as Jaffrain et al. (2011). Filtered disdrometer data were then used to estimate rainfall intensity, and equivalent attenuation at 22.715 GHz was calculated using the T-matrix approach developed by Mishchenko and Travis (1994) using below equations:

$$R = 6\pi \times 10^{4} \int_0^{\infty} D^3 v(D) N(D) dD \tag{1}$$

$$k = \frac{1}{\ln(10)} \int_0^{\infty} \sigma_E(D) N(D) dD \tag{2}$$

where $v(D)$ is the raindrop terminal fall speed (m $s^{-1}$) as a function of the equivalent spherical drop diameter $D$ (mm), $N(D)dD$ ($m^{-3}$) is the total number of drops in the diameter interval ($D$, $D + dD$) per unit volume, $R$ is the rainfall rate (mm $hr^{-1}$), $k$ is the specific attenuation (dB $km^{-1}$), and $\sigma_{ext}$ is extinction cross section ($cm^2$) of a hydrometeor with a diameter $D$ (mm).

### 2.2. Commercial Microwave Link Data

For this study, received signal level (RSL) data from one Motorola commercial microwave link (CML) were collected, with an OTT PARSIVEL[1] disdrometer installed at one end. This link operated at a frequency of 22.715 GHz with a path length of 3.79 km and a constant transmit power over time. The signal was sampled at 10 Hz with a resolution of 0.1 dB, but only minimum, maximum, and average received powers over a 15-min interval were stored for monitoring the quality of the network. These RSL data were later used to estimate the minimum, maximum, and average attenuation using the algorithm presented in Overeem et al. (2016a).

## 3. Methodology

### 3.1. Long Short Term Memory Network

Long short-term memory (LSTM) architecture is a specific type of recurrent neural network (RNN) which was originally designed to capture long-term dependencies in time series data. It has the capability of overcoming issues of vanishing and exploding gradients (Hochreiter & Schmidhuber, 1997). This architecture preserves the states over a longer period of time without losing temporal dependencies (Hochreiter & Schmidhuber, 1997). For the problem related to the temporal distribution with non-linearity in data, such as natural language processing, image classification, and sound translation, this method has proved the most useful when compared with other statistical and conventional feed-forward models (Shen, 2018; Shen et al., 2018).

In this study, among various architectures within the LSTM, the sequence-to-sequence LSTM network has been used herein (Figure 1a). The output $y = y_1, y_2, ..., y_n$ is the average attenuation from an input $x = x_1, x_2, ..., x_n$ consisting of $n$ consecutive time steps. The input variables are the minimum and maximum attenuation, and the output is the average attenuation.



Printed by [Monash University - 149.171.191.252 - /doi/epdf/10.1029/2019WR026255] at [24/09/2020].

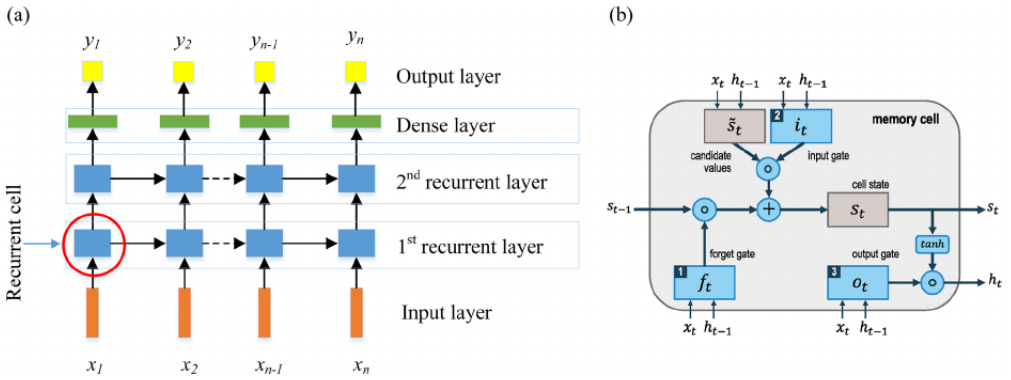


**Figure 1.** (a) General architecture of a two-layer sequence-to-sequence recurrent neural network. The output of the second layer for each of the time steps is fed into a dense layer to calculate the prediction ($y_1$, $y_2$..., $y_n$). (b) The internal architecture of the LSTM recurrent cell (adapted from Hu et al., 2018).

In each time step $t$ ($1 \leq t \leq n$), the current input $x_t$ is processed in the LSTM recurrent cells of each layer in a network as in Figure 1b. The LSTM cell is composed of an input layer, one or more memory cells, and an output layer. The major feature of LSTM networks is that they contain hidden layers, which are referred to as memory cells. Each of these memory cells is composed of three gates for adjusting the internal cell state ($s_t$): the forget gate ($f_t$), an input gate ($i_t$), and an output gate ($o_t$). Details of the LSTM algorithm are explained by Hochreiter and Schmidhuber (1997) and can be summarized as

$$g_t = tanh\left(W_{gx}x_t + W_{gh}h_{(t\ \ 1)} + b_g\right) \tag{3}$$

$$i_t = \sigma\left(W_{ix}x_t + W_{ih}h_{(t\ \ 1)} + b_i\right) \tag{4}$$

$$f_t = \sigma\left(W_{fx}x_t + W_{fh}h_{(t\ \ 1)} + b_f\right) \tag{5}$$

$$o_t = \sigma\left(W_{ox}x_t + W_{oh}h_{(t\ \ 1)} + b_0\right) \tag{6}$$

$$s_t = g_t \quad i_t + s_{t\ \ 1} \quad o_t \tag{7}$$

$$h_t = \tanh(s_t) \quad o_t \tag{8}$$

$$y_t = \left(W_{hy}h_t + b_y\right) \tag{9}$$

where $\sigma$ is the logistic sigmoidal function, is an element wise multiplication, $g_t$ is input node, $x_t$ is the input forcing, $W$ are the network weights, $b$ are bias parameters, and $y$ is the output for the time step $t$. Among these parameters, $W$ and $b$ are learnable (adaptable) parameters which are updated at each time step based on a given loss function.

### 3.2. Modeling Approach

For this study, a two-layer LSTM model was designed with fully connected hidden layers, 50 LSTM neurons, and a single dense layer for predicting output was implemented to form the base architecture of the model. This architecture was adopted based on the prior sensitivity analysis using different numbers of layers and neurons. The mean squared error (MSE) was used as a loss function (quantified and minimized for average attenuation) and the adaptive moment estimation, in short known as Adam, was used to optimize the model. Adam is an adaptive learning rate optimization algorithm proposed by Diederik and Ba (2014) which performs better than the conventional stochastic gradient descent algorithms.

This model was run for 200 epochs (i.e., 200 passes through the data), with variations in the number of layers and hidden neurons tested to understand the sensitivity and optimize model performance. A newly developed architecture of the RNN called grated recurrent unit (GRU), developed by Cho et al. (2014), was also tested. The GRU has the similar goal of tracking the long-term dependencies in the time series data. This GRU contains two

Printed by [Monash University - 049.127.091.232 - /doi/epdf/10.1029/2019WR026255] at [24/09/2020].

gates called a reset gate and an update gate as opposed to the LSTM, which contains three gates. Similarly, a conventional artificial neural network (which is a memoryless model that does not capture the temporal trend in the data) using dense layers was tested and compared to the performance of the LSTM model. In order to prevent the model from overfitting, a dropout rate of 50% was adopted for each of the layers.

The input data set (15-min maximum and minimum attenuation) for the model was produced from the 30-s disdrometer-derived attenuation. This whole data set was split into sets of sub-sequences of size $N_w$ which helps in back-propagation of the LSTM model over time. This is helpful in back-propagation through time, which unfolds this LSTM into a feed-forward neural network with a number of recurrent steps, $N_w$. This $N_w$ was determined based on the most prevalent duration of rain events (based on the Poisson-type distribution). Time steps of 10, 15, and 20 min were considered for this. Samples with all dry time steps were discarded. For the rain events longer than the considered time steps, these events were split into overlapping sequences. Here we define a rainfall event as a rain period separated by a 1-hr or longer rain-free period, and having a minimum rainfall rate of 0.1 mm $hr^{-1}$.

The chosen data set was split into two groups of sequences: (i) the training group corresponding to 80% of the whole sequence and (ii) the testing group corresponding to the remaining 20% of the whole sequence (typical independent test amount range between 10% and 20%). And within a training group, 80% of the data was used to optimize the model, and the remaining 20% was used as a validation set to monitor the learning process. This validation was conducted to make sure that there was no overfitting of the model parameters during the training phase.

### 3.3. Rainfall Retrieval From CML Data

The algorithm introduced by Overeem et al. (2016a) was used for the rainfall retrieval from commercial microwave link data. A brief description of the algorithm is given below:

1. Wet/dry classification: The algorithm proposed by Overeem et al. (2016a) uses a spatial correlation by looking at the nearby links to identify the given time steps as either wet or dry. In the case presented herein, there was only one link, and so this classification was based solely on the disdrometer data. Time steps for which the rainfall rate observed by the disdrometer was greater than or equal to 0.1 mm $hr^{-1}$ were classified as "wet," and all remaining time steps were classified as "dry."
2. Identification of the reference/baseline signal: This was calculated as the moving median of the signal level during the previous 24-hr dry period.
3. Wet antenna attenuation correction: A constant wet antenna attenuation was obtained based on the optimization a microwave link using the algorithm suggested by de Vos et al. (2019). In this case, average RSL data were used to obtain the wet antenna attenuation of 1.2 dB. The resultant (total wet antenna) attenuation was later divided by the path length to obtain the rain-induced specific attenuation ($k$).
4. Computation of the rainfall rate: This was done using the power law relationship between rainfall intensity ($R$) and specific attenuation ($k$) (Olsen et al., 1978) as

$$R = ak^b \tag{10}$$

where $R$ is the rainfall intensity (mm $hr^{-1}$), $k$ is the specific attenuation of the signal (dB $km^{-1}$), and $a$ and $b$ are parameters depending on the frequency, polarization, drop size distribution, drop shape, and canting angle. The values for the parameters $a$ and $b$ ($a = 9.563$ and $b = 0.956$) were derived for Melbourne for an equivalent 22.715 GHz microwave link using data obtained from an OTT PARSIVEL1 optical disdrometer over a 3-year period (Guyot et al., 2019).

## 4. Results

Figure 2 shows an illustrative comparison of the rainfall estimation (obtained from the attenuation estimate using Equation 10 with the same $a$ and $b$ parameters) using the weighted average method with parameter $\alpha = 0.21$ and the LSTM model. The factor $\alpha = 0.21$ was obtained by optimizing the minimum and maximum attenuation values. It shows how the LSTM model outperformed the constant weighted average method when compared with observed attenuation and rainfall directly derived from the 30-s disdrometer data.



Printed by [Monash University - 149.171.091.252 - /doi/epdf/10.1029/2019WR026255] at [24/09/2020].

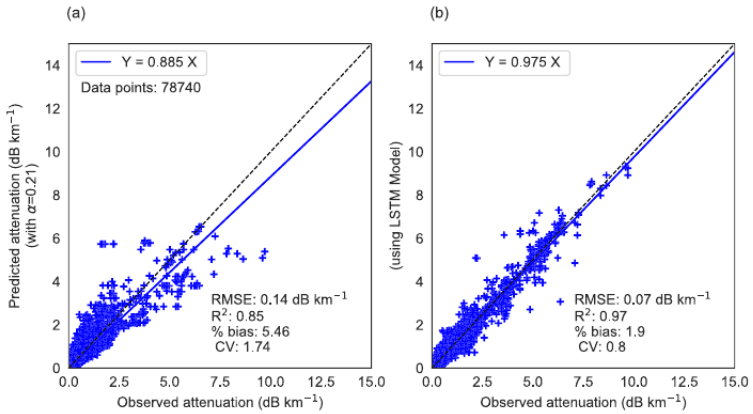


**Figure 2.** Comparison of the predicted to the observed (a, b) average attenuation based on disdrometer-derived attenuation and (c, d) rainfall rate based on disdrometer-derived attenuation estimates for the weighted average method with = 0.21 and for the LSTM model.

Using the weights 21% and 79% for the maximum and minimum, respectively, resulted in a larger relative bias in both average attenuation and rainfall rate compared to using the LSTM model. The root mean square error (RMSE) between the attenuation measured by the disdrometer and that obtained from LSTM dropped from 0.14 to 0.07 dB km$^{-1}$, while the relative bias reduced from 5.6% to 1.9% and the coefficient of determination ($R^2$) increased from 0.85 to 0.97. Similar improvements were observed for rainfall intensity. To further examine the performance of the LSTM model over a constant weighted average method, other values, ranging from 0 to 1 with a step of 0.1, were also tested. It was found that the statistics (RMSE, $R^2$, CV, and percentage bias) of the LSTM model were better compared to those of the constant weighted average method for all values of considered.

### 4.1. Performance of Various Architectures of the Deep Learning Model

Table 1 shows five statistical measures for three different architectures of the deep learning model. Among these, the dense layer was a traditional artificial neural network with no capacity to model time dependencies, and other two (GRU and LSTM) were different types of RNN architectures. For both the training and test data sets, the performance of the LSTM and GRU were very similar, while the GRU performed better in terms of relative bias. The remaining four statistical parameters were similar. The RMSE, MAE, $R^2$, and coefficient of variation (CV) values for the LSTM model indicate that it performed slightly better than the GRU and much better than the dense layer.

### 4.2. Sensitivity of the Model

Figure 3 shows the LSTM model sensitivity for different numbers of neurons and hidden layers for the training and test data sets. The RMSE decreased significantly with an increasing amount of neurons up to 100 and

**Table 1**
*Comparison of the Performance of Various Model Architectures for Rainfall Prediction for the Training and Test Data Sets Based on the Disdrometer*

| Data sets | Architecture | $R^2$ | RMSE (mm hr$^{-1}$) | MAE (mm hr$^{-1}$) | Relative bias (%) | CV |
|---|---|---|---|---|---|---|
| **Training data set** | | | | | | |
| 31,495 sets | Dense layer (ANN) | 0.85 | 1.12 | 0.21 | 6.23 | 1.48 |
| | GRU | 0.96 | 0.61 | 0.17 | 0.17 | 0.81 |
| | LSTM | 0.97 | 0.62 | 0.17 | 2.87 | 0.76 |
| **Test data set** | | | | | | |
| 7,874 sets | Dense layer (ANN) | 0.86 | 1.11 | 0.20 | 0.13 | 1.47 |
| | GRU | 0.95 | 0.71 | 0.20 | 0.45 | 0.84 |
| | LSTM | 0.97 | 0.64 | 0.19 | 2.86 | 0.83 |

Printed by [Monash University - 149.171.091.252 - /doi/epdf/10.1029/2019WR026255] at [24/09/2020].

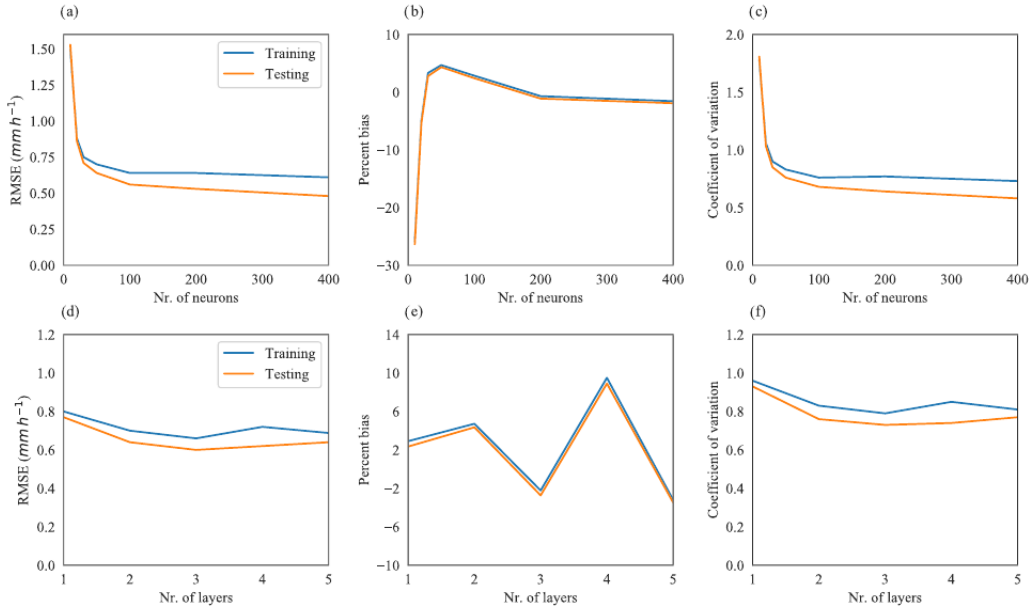


**Figure 3.** LSTM model performance for (a) RMSE; (b) relative percentage bias; (c) coefficient of variation versus number of hidden neurons; (d) RMSE; (e) relative percentage bias; and (f) oefficient of variation versus number of hidden layers.

remained constant after increasing the number of neurons further. Similar characteristics were observed for the relative bias and the coefficient of variation. For the hidden layers, the LSTM model performance increased from one to three layers and started to deteriorate beyond that value. The performance of the model showed the fluctuation based on percent bias, but its best performance was observed with three hidden layers.

### 4.3. Application of the Disdrometer Trained LSTM Model to the CML Data Set

Figure 4 shows the comparison of the predicted rainfall using the CML observations for the 22 GHz CML link closest to the disdrometer site. The factor = 0.30 and wet antenna attenuation of 1.2 dB were obtained

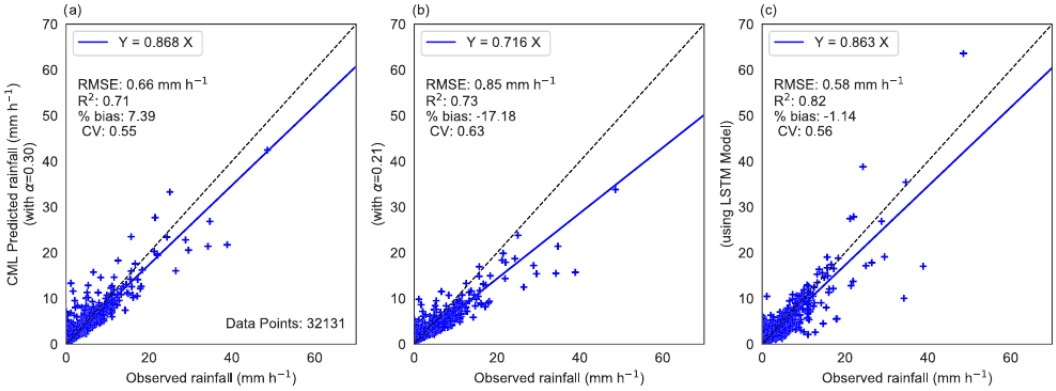


**Figure 4.** Scatter plots of the predicted versus observed rainfall intensity using commercial microwave link data for (a) the weighted average method with = 0.30; (b) the weighted average method with = 0.21; and (c) the disdrometer-trained LSTM model.



Printed by [Monash University - 149.171.091.252 - /doi/epdf/10.1029/2019WR026255] at [24/09/2020].

through optimization of average rainfall obtained from minimum and maximum rainfall based on the gauge-adjusted radar data obtained from the Bureau of Meteorology. The LSTM model outperformed the two weightage average method results ( = 0.30 and 0.21) for all but a few events. In particular, shorter duration convective rainfall events were under-estimated by the LSTM model compared with the weighted average method.

## 5. Discussion and Conclusion

This paper demonstrated a novel approach for improving the rainfall estimation from commercial microwave links (when only limited information such as minimum and maximum RSL data are available) by using a deep learning model. Results showed a significant improvement based on simulated microwave link attenuation data from a disdrometer and for real commercial microwave link data, compared with two weighted average method with parameters using = 0.31 or = 0.21. Although the performance of the deep learning model was lower, using the commercial microwave link data than using the simulated data, there was still a good improvement in bias and $R^2$ of the rainfall estimation compared to the weightage average method. The reduction in performance can be due to number of factors, linked to the fact that the type of the data derived from the CML is different from the attenuation data derived from the disdrometer: (1) CML data are derived from a sampling of 10 Hz and the corresponding minimum and maximum within each 15-min period, as opposed to the accumulation over 30 s for the disdrometer. (2) The attenuation from the CML is an integration over a path length of 3.8 km as opposed to being a point measurement for the disdrometer. (3) In order to derive the attenuation from the CML data, the baseline received signal level had to be subtracted from the RSL at time $t$, and this baseline varies between events and thus induces some uncertainties in the data. Additionally, attenuation due to a wet antenna plays a crucial role in the rainfall retrieval from CML data. This wet antenna effect strongly depends on the material of the antenna cover used for the transmitter and receiver (van Leth et al., 2018). Attenuation as simulated using disdrometer data does not include this wet antenna effect, which therefore represents one limitation of our proposed method. Wet antenna attenuation is usually dependent on the rainfall intensity (Schleiss et al., 2013), thus not considering this phenomenon further adds uncertainty to the model performance. This was also observed based on the analysis using additional bias (which is equivalent to the effect of wet antenna attenuation). In addition, quantization of the raw data also impacted on the rainfall retrieval. However, for the frequency and path length of the CML used in this study, this impact was negligible when compared to the effect of wet antenna attenuation.

This paper only shows results based on a CML link that is close to the disdrometer, but a similar methodology could be adopted for other links having different frequencies within the same climate. Usually links having frequencies ranging from 20 to 40 GHz could be more suitable as these tend to be associated with a close to linear specific attenuation-rain rate relationship (Berne & Uijlenhoet, 2007). Following our approach and based on time series of attenuation and rainfall rates from a single disdrometer, various deep learning models for each given frequency can be designed and implemented. The limitations of this data-driven approach reside first in the size and representativeness of the collected disdrometer data set for a given location. In this work, we have demonstrated the feasibility of using a disdrometer-trained LSTM model to predict rainfall for a nearby link. The applicability of this trained model to retrieve rainfall for a wider area still has to be demonstrated, and likely the size (duration, diversity, and quantity of recorded rainfall events) of the disdrometer data will be an important factor. Secondly, this disdrometer-trained model replaces only one of the steps among numerous successive steps in the rainfall retrieval, such as dry/wet classification, baseline estimation, and wet antenna attenuation. Such data-driven deep learning approaches for dry/wet classification have also been recently explored both by Polz et al. (2019) (under review at HESSD) and Habi and Messer (2018). Similarly, one could think of a similar approach for baseline estimation and wet antenna attenuation. Finally, adaptive learning and/or transfer learning techniques could also be implemented as additional steps to improve the overall rainfall estimates from CML data.

### Data Availability Statement

The data set presented in this study are publicly available online (https://doi.org/10.5281/zenodo.3629929). These data include raw data from OTT[1] disdrometer and received signal level data (minimum, maximum,

Printed by [Monash University - 049.17.091.252 - /doi/epdf/10.1029/2019WR026255] at [24/09/2020].

and average) from a commercial microwave link. Detail information of the data set can be found on description file.

**Acknowledgments**
The authors would like to thank Melbourne Water for providing the site for installation of disdrometer and Motorola for providing received signal level data for the microwave link close to the disdrometer. This work is supported by an Australian Research Council Discovery grant number DP160101377. Jayaram Pudashine also acknowledges the top-up support from CSIRO Data 61for his PhD studies.

## References

Berne, A., & Uijlenhoet, R. (2007). Path-averaged rainfall estimation using microwave links: Uncertainty due to spatial rainfall variability. *Geophysical Research Letters*, *34*, L07403. https://doi.org/10.1029/2007GL029409

Bianchi, B., Rieckermann, J., & Berne, A. (2013). Quality control of rain gauge measurements using telecommunication microwave links. *Journal of Hydrology*, *492*, 15–23. https://doi.org/10.1016/j.jhydrol.2013.03.042

Brauer, C., Overeem, A., Leijnse, H., & Uijlenhoet, R. (2016). The effect of differences between rainfall measurement techniques on groundwater and discharge simulations in a lowland catchment. *Hydrological Processes*, *30*(21), 3885–3900. https://doi.org/10.1002/hyp.10898

Cho, K., van Merrienboer, B., Bahdanau, D., & Bengio, Y. (2014). On the properties of neural machine translation: Encoder-decoder approaches. *arXiv e prints*. https://ui.adsabs.harvard.edu/abs/2014arXiv1409.1259C

Chwala, C., Gmeiner, A., Qiu, W., Hipp, S., Nienaber, D., Siart, U., et al. (2012). Precipitation observation using microwave backhaul links in the alpine and pre-alpine region of southern Germany. *Hydrology and Earth System Sciences*, *16*(8), 2647–2661. https://doi.org/10.5194/hess-16-2647-2012

Chwala, C., & Kunstmann, H. (2019). Commercial microwave link networks for rainfall observation: Assessment of the current status and future challenges. *Wiley Interdisciplinary Reviews Water*, *6*(2), e1337. https://doi.org/10.1002/wat2.1337

de Vos, L. W., Overeem, A., Leijnse, H., & Uijlenhoet, R. (2019). Rainfall estimation accuracy of a nationwide instantaneously sampling commercial microwave link network: Error dependency on known characteristics. *Journal of Atmospheric and Oceanic Technology*, *36*(7), 1267–1283. https://doi.org/10.1175/jtech-d-18-0197.1

Diederik, K., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv e prints*, Retrieved from https://arxiv.org/abs/1412.6980v9

Eshel, A., Messer, H., Ostrometzky, J., Raich, R., Alpert, P., & Laronne, J. B. (2017). On the use of measurements from a commercial microwave link for evaluation of flash floods in arid regions. *Atmospheric Chemistry and Physics Discussions*, 1–25. https://doi.org/10.5194/acp-2017-963

Fencl, M., Rieckermann, J., Schleiss, M., Stransky, D., & Bares, V. (2013). Assessing the potential of using telecommunication microwave links in urban drainage modelling. *Water Science and Technology*, *68*(8), 1810–1818. https://doi.org/10.2166/wst.2013.429

Gazit, L., & Messer, H. (2018). Sufficient conditions for reconstructing 2-D rainfall maps. *IEEE Transactions on Geoscience and Remote Sensing*, *56*(11), 1–10. https://doi.org/10.1109/tgrs.2018.2836998

Goldshtein, O., Messer, H., & Zinevich, A. (2009). Rain rate estimation using measurements from commercial telecommunications links. *IEEE Transactions on Signal Processing*, *57*(4), 1616–1625. https://doi.org/10.1109/TSP.2009.2012554

Gosset, M., Kunstmann, H., Zougmore, F., Cazenave, F., Leijnse, H., Uijlenhoet, R., et al. (2016). Improving rainfall measurement in gauge poor regions thanks to mobile telecommunication networks. *Bulletin of the American Meteorological Society*, *97*(3), ES49–ES51. https://doi.org/10.1175/bams-d-15-00164.1

Graf, M., Chwala, C., Polz, J., & Kunstmann, H. (2019). Rainfall estimation from a German-wide commercial microwave link network: Optimized processing and validation for one year of data. *Hydrology and Earth System Sciences Discussions*. https://doi.org/10.5194/hess-2019-423

Graves, A., & Schmidhuber, J. (2009). Offline handrwriting recognition with multidimensional recurrent neural network. *Advances in Neural Information Processing Systems*, 545–552.

Guyot, A., Pudashine, J., Protat, A., Uijlenhoet, R., Pauwels, V. R. N., Seed, A., & Walker, J. P. (2019). Effect of disdrometer type on rain drop size distribution characterisation: a new dataset for south-eastern Australia. *Hydrology and Earth System Sciences*, *23*(11), 4737–4761. https://doi.org/10.5194/hess-23-4737-2019

Habi, H. V., & Messer, H. (2018). Wet-Dry Classification using LSTM and Commerical Microwave Links, IEEE 10th Sensory Array and Multichannel Signal Processing Workshop, 149-153. https://doi.org/10.1109/SAM.2018.8448679

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.

Hu, C., Wu, Q., Li, H., Jian, S., Li, N., & Lou, Z. (2018). Deep learning with a long short-term memory networks approach for rainfall-runoff simulation. *Water*, *10*(11), 1543. https://doi.org/10.3390/w10111543

Jaffrain, J., & Berne, A. (2011). Experimental quantification of the sampling uncertainty associated with measurements from PARSIVEL disdrometers. *Journal of Hydrometeorology*, *12*(3), 352–370. https://doi.org/10.1175/2010jhm1244.1

Jaffrain, J., Studzinski, A., & Berne, A. (2011). A network of disdrometers to quantify the small-scale variability of the raindrop size distribution. *Water Resources Research*, *47*, W00H06. https://doi.org/10.1029/2010WR009872

Kavetski, D., Fenicia, F., & Clark, M. P. (2011). Impact of temporal data resolution on parameter inference and model identification in conceptual hydrological modeling: Insights from an experimental catchment. *Water Resources Research*, *47*, W05501. https://doi.org/10.1029/2010WR009525

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall–runoff modelling using long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences*, *22*(11), 6005–6022. https://doi.org/10.5194/hess-22-6005-2018

Leijnse, H., Uijlenhoet, R., & Stricker, J. N. M. (2007). Rainfall measurement using radio links from cellular communication networks. *Water Resources Research*, *43*, W03201. https://doi.org/10.1029/2006WR005631

Leijnse, H., Uijlenhoet, R., & Stricker, J. N. M. (2008). Microwave link rainfall estimation: Effects of link length and frequency, temporal sampling, power resolution, and wet antenna attenuation. *Advances in Water Resources*, *31*(11), 1481–1493. https://doi.org/10.1016/j.advwatres.2008.03.004

Liberman, Y., Samuels, R., Alpert, P., & Messer, H. (2014). New algorithm for integration between wireless microwave sensor network and radar for improved rainfall measurement and mapping. *Atmospheric Measurement Techniques*, *7*(10), 3549–3563. https://doi.org/10.5194/amt-7-3549-2014

Loffler-Mang, M., & Joss, J. (2000). An optical disdrometer for measuring size and velocity of hydrometers. *Journal of Atmospheric and Oceanic Technology*, *2*, 130–139. https://doi.org/10.1175/1520-0426(2000)017<0130:aodfms>2.0.co;2

Mathieu, M., Couprie, C., & LeCun, Y. (2015). Deep multi-scale video prediction beyond mean square error. *arXiv e prints*. https://ui.adsabs.harvard.edu/abs/2015arXiv151105440M

Messer, H. (2018). Capitalizing on cellular technology—Opportunities and challenges for near ground weather monitoring †. *Environments*, *5*(7), 73. https://doi.org/10.3390/environments5070073

Printed by [Monash University - 149.171.091.252 - /doi/epdf/10.1029/2019WR026255] at [24/09/2020].

Messer, H., Zinevich, A., & Alpert, P. (2006). Environmental monitoring by wireless communication networks. *Science*, *312*(5774), 713–713. https://doi.org/10.1126/science.1120034

Mishchenko, M. I., & Travis, L. D. (1994). T-matrix computations of light scattering by large spheroidal particles. *Optics Communications*, *109*(1), 16–21. https://doi.org/10.1016/0030-4018(94)90731-5

Mishra, V. K., Gharanjik, A., Bhavani, M. R., & Bjorn, O. (2018). Deep learning framework for precipitation retrievals from communication satellites. Paper Presented at the 10th European Conference on Radar in Meteorology and Hydrology, Wageningen, Netherlands.

Olsen, R., Rogers, D. V., & Hodge, D. B. (1978). The aRb relation in the calculation of rain attenuation. *IEEE Transactions on Antennas and Propagation*, *26*(2), 318–329. https://doi.org/10.1109/TAP.1978.1141845

Ostrometzky, J., & Messer, H. (2014). Accumulated rainfall estimation using maximum attenuation of microwave radio signal, 2014 IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM), A Coruna, 193–196. https://doi.org/10.1109/SAM.2014.6882373

Ostrometzky, J., & Messer, H. (2017). Comparison of different methodologies of parameter-estimation from extreme values. *IEEE Signal Processing Letters*, *24*(9), 1293–1297. https://doi.org/10.1109/lsp.2017.2723544

Overeem, A., Leijnse, H., & Uijlenhoet, R. (2011). Measuring urban rainfall using microwave links from commercial cellular communication networks. *Water Resources Research*, *47*, W12505. https://doi.org/10.1029/2010WR010350

Overeem, A., Leijnse, H., & Uijlenhoet, R. (2013). Country-wide rainfall maps from cellular communication networks. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(8), 2741–2745. https://doi.org/10.1073/pnas.1217961110

Overeem, A., Leijnse, H., & Uijlenhoet, R. (2016a). Retrieval algorithm for rainfall mapping from microwave links in a cellular communication network. *Atmospheric Measurement Techniques*, *9*(5), 2425–2444. https://doi.org/10.5194/amt-9-2425-2016

Overeem, A., Leijnse, H., & Uijlenhoet, R. (2016b). Two and a half years of country-wide rainfall maps using radio links from commercial cellular telecommunication networks. *Water Resources Research*, 8039–8065. https://doi.org/10.1002/2016WR019412

Polz, J., Chwala, C., Graf, M., & Kunstmann, H. (2019). Rain event detection in commercial microwave link attenuation data using convolutional neural networks. *Atmospheric Measurement Techniques Discussions*, *2019*, 1–22. https://doi.org/10.5194/amt-2019-412

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat (2019). Deep learning and process understanding for data-driven earth system science. *Nature*, *566*(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1

Rios Gaona, M. F., Overeem, A., Raupach, T. H., Leijnse, H., & Uijlenhoet, R. (2018). Rainfall retrieval with commercial microwave links in São Paulo, Brazil. *Atmospheric Measurement Techniques*, *11*(7), 4465–4476. https://doi.org/10.5194/amt-11-4465-2018

Schleiss, M., Rieckermann, J., & Berne, A. (2013). Quantification and modeling of wet-antenna attenuation for commercial microwave links. *IEEE Geoscience and Remote Sensing Letters*, *10*(5), 1195–1199. https://doi.org/10.1109/LGRS.2012.2236074

Shen, C. (2018). A transdisciplinary review of deep learning research and its relevance for water resources scientists. *Water Resources Research*, *54*, 8558–8593. https://doi.org/10.1029/2018WR022643

Shen, C., Laloy, E., Elshorbagy, A., Albert, A., Bales, J., Chang, F. J., et al. (2018). HESS opinions: Incubating deep-learning-powered hydrologic science advances as a community. *Hydrology and Earth System Sciences*, *22*(11), 5639–5656. https://doi.org/10.5194/hess-22-5639-2018

Shi, X., Gao, Z., Lausen, L., Wang, H., & Yeung, D. Y. (2014). Deep learning for precipitation nowcasting: A benchmark and a new model. *Advances in Neural Information Processing Systems*, *30*, 5617–5627.

Smiatek, G., Keis, F., Chwala, C., Fersch, B., & Kunstmann, H. (2017). Potential of commercial microwave link network derived rainfall for river runoff simulations. *Environmental Research Letters*, *12*(3), 034026. https://doi.org/10.1088/1748-9326/aa5f46

Sohail Afzal, M., Shah, S. H. M., Cheema, M. J. M., & Ahmad, R. (2018). Real time rainfall estimation using microwave signals of cellular communication networks: A case study of Faisalabad, Pakistan. *Hydrology and Earth System Sciences Discussions*, 1–20. https://doi.org/10.5194/hess-2017-740

Uijlenhoet, R., Overeem, A., & Leijnse, H. (2018). Opportunistic remote sensing of rainfall using microwave links from cellular communication networks. *Wiley Interdisciplinary Reviews Water*, *5*(4), e1289. https://doi.org/10.1002/wat2.1289

van Leth, T. C., Overeem, A., Leijnse, H., & Uijlenhoet, R. (2018). A measurement campaign to assess sources of error in microwave link rainfall estimation. *Atmospheric Measurement Techniques*, *11*(8), 4645–4669. https://doi.org/10.5194/amt-11-4645-2018

Printed by [Monash University - 149.171.091.252 - /doi/epdf/10.1029/2019WR026255] at [24/09/2020].