Check for updates

High temporal resolution rainfall rate estimation from rain gauge measurements

Yang Song, Dawei Han and Miguel A. Rico-Ramirez

ABSTRACT

Rainfall rates derived from tipping bucket rain gauges generally ignore the detailed variation at a finer temporal scale that particularly occurs in light rainfall events. This study extends the exploration of using artificial neural networks (ANNs), in comparison with the conventional linear interpolation method (LIM) and the cubic spline algorithm (CSA) for rainfall rate estimation at fine temporal resolution using rain gauge data based on a case study at Chilbolton and Sparsholt observatories, UK. A supervised feed-forward neural network integrated with the backpropagation algorithm is used to identify the complex nonlinear relationships between input and target variables. The results indicate that the ANN considerably outperforms the CSA and LIM with higher Nash–Sutcliffe efficiency, lower root mean square error and lower rainfall amount differences when compared to the disdrometer observations when the model is trained within a broad span of input values. Consistent stability in accurately estimating rainfall rate in different sites shows the intrinsic advantage of ANNs in learning and self-adaptive abilities in modelling complex nonlinear relationships between the inputs and target variables.

Key words | ANN, CSA, disdrometer, LIM, rainfall rate/intensity, tipping bucket rain gauge

Yang Song (corresponding author)
Dawei Han
Miguel A. Rico-Ramirez
Department of Civil Engineering,
University of Bristol,
Bristol BS8 1TR,
UK
E-mail: *yang.song@bristol.ac.uk*

INTRODUCTION

Precipitation is a fundamental input for practically all aspects of hydrological assessment and modelling. However, it is usually exceptionally hard to derive precise measurements (Savina *et al.* 2012). Both rain gauges and weather radar are commonly used to monitor precipitation (Sideris *et al.* 2014), but the most widely used instruments for point measurements of precipitation of ground truth are rain gauges (Liu *et al.* 2013), especially with tipping bucket rain gauges (TBRs) due to their low cost and high reliability. For the estimation of areal precipitation, weather radars are broadly applied at both high temporal and spatial resolutions in the hydrological and meteorological communities (Rico-Ramirez *et al.* 2012). Although previous studies (Garcia-Pintado *et al.* 2009; Kirstetter *et al.* 2015) have been conducted in merging and processing rainfall products by combining rain gauges and weather radars, constraints still exist in estimating the 'true rainfall' due to the spatial and temporal differences between them.

Due to point sampling characteristics of rain gauges, typically, rain gauge data are considered to provide high accuracy of point measurement on the ground level. However, rain gauges are generally distributed sparsely, therefore resulting in inadequacies of capturing the spatial variability of precipitation in space (Savina *et al.* 2012). On the other hand, radars provide a spatially dense and wide coverage of areal precipitation indirectly in elevated volume, but with uncertainties on quantitative precipitation estimation on the ground (Sideris *et al.* 2014). As a result, discrepancies can be found between radars and rain gauges due to the huge magnitude differences in spatial resolution as

doi: 10.2166/hydro.2017.054

Downloaded from http://iwaponline.com/jh/article-pdf/19/6/930/659024/jh0190930.pdf
by guest

well as inherent errors in both radar and gauge measurements.

Instantaneous precipitation measurements can be derived from weather radars at high temporal resolutions typically ranging from 5 to 10 minutes and modern X-band radars can even provide less than 1 minute temporal resolution. Weather radars transmit electromagnetic waves in the microwave range (usually 3–10 GHz) into the atmosphere and a small percentage of energy is reflected back to the radar antenna when precipitation particles (also known as hydrometeors) exist along the direction of the transmitted waves. The reflected energy (or reflected power) can be related to the radar reflectivity Z, which in turn, can be related to the rainfall intensity R by applying a standard nonlinear Z–R relationship (Garcia-Pintado *et al.* 2009). In comparison, rain gauges can constantly collect and accumulate rainfall over a time period of interest (Sideris *et al.* 2014). A linear interpolation method (LIM) which simply divides the tipping bucket volume by the time between tips is applied to obtain the mean rainfall rates. However, the mean rainfall rate derived from TBRs normally do not consider the variations of precipitation for rainfall events with short duration (Habib *et al.* 2001). Since the high temporal resolution of rainfall measurements can significantly eliminate the temporal mismatch in precipitation measurements between weather radar and TBRs, improvement to radar rainfall adjustment could be achieved by reducing the temporal representation uncertainty of TBR data. It is therefore crucial to estimate which heavy rainfall intensities derived from TBR can either underestimate or overestimate the true rainfall rates (as shown later for group 3 and 4, respectively), that can significantly affect the hydrological response of urban catchments during flash flooding events (Ochoa-Rodriguez *et al.* 2015). Therefore, it is desirable that rainfall rates at a finer temporal resolution should be obtained from TBR measurements.

Many prior studies have attempted to use TBRs to retrieve high temporal resolution rainfall rates (Sadler & Busscher 1989; Habib *et al.* 2001; Wang *et al.* 2008). Sadler & Busscher (1989) and Wang *et al.* (2008) have produced rainfall rates at 1-minute temporal resolution using the cubic spline algorithm (CSA) by fitting the accumulated rainfall amount from TBRs, and then using the derivative of the cubic spline to compute the high temporal resolution rainfall rates. Compared with traditional methods such as linear or quadratic approaches, the CSA can be easily implemented (Sadler & Busscher 1989); moreover, a seamless fitting curve can only be generated by applying a cubic or higher order algorithm (Wang *et al.* 2008). Nevertheless, the CSA requires the piecewise continuous property, and it may not be the best choice for rainfall data that are responsive to the smoothness of third or higher order derivatives. Furthermore, it may not be suitable for heavy rainfall events with short duration (e.g., less than 2 minutes) where either a first or second derivative cannot be determined from precipitation information. In addition, large rainfall gradients at low rainfall rates can produce negative rainfall rates when applying the CSA (Wang *et al.* 2008).

The raindrop size distributions and the fall velocities of precipitation particles (also known as hydrometeors) at high temporal resolutions can be measured by using a disdrometer, which can be used to estimate rainfall intensity, radar reflectivity, water content and other radar measurements (Sieck *et al.* 2007). Disdrometers can also help to detect some of the sources of errors in radar rainfall estimations (Islam *et al.* 2012). As a result, the disdrometer has been extensively adopted in rain gauge, radar and satellite-borne remote sensing research (Bringi & Chandrasekar 2001; Wang *et al.* 2008; Islam *et al.* 2012; Liu *et al.* 2013). For the measurement of drop size distributions as well as its quantification at the ground, the Joss–Waldvogel disdrometer (JWD) is traditionally used as the reference instrument among all different types of disdrometers (Tokay *et al.* 2005). Besides JWD, the 2-D video disdrometer (Thurai & Bringi 2005) and the laser optical disdrometer (Jaffrain & Berne 2011) are also widely used and can directly measure rain drop sizes as well as classify the hydrometeor types based on characteristics such as rain drop sizes, shapes and fall velocities (Liu *et al.* 2013).

The objective of this study is to retrieve high temporal resolution rainfall information by combining TBRs and JWD measurements. Recent advances in the applications of artificial neural network (ANN) approaches which, instead of using existing numerical equations, integrate data of rain drop size distributions derived from disdrometer, radar reflectivity and water content to simulate rainfall rates through a learning process that minimizes the errors between theoretical and experimental outputs (Tengeleng & Armand 2014).

Downloaded from http://iwaponline.com/jh/article-pdf/19/6/930/659024/jh0190930.pdf
by guest

Such research has proved that ANNs are capable of rainfall rate estimation; however, there is a research gap in retrieving high temporal resolution rainfall information from a combination of rainfall measurements (e.g., TBRs and disdrometers). Moreover, the estimated rainfall rates are not verified with those derived from any other rain gauges, especially for the TBRs which are commonly considered as the ground truth (Sieck *et al.* 2007). Nevertheless, the ANN is a heuristic and prospective method, which could be utilized in retrieving rainfall rates at high temporal resolution based on data from TBRs, as ANN can implicitly identify complex nonlinear relationships between input and target values. Therefore, an ANN model integrated with 1-minute rainfall rate is set up and assessed with a TBR and a JWD. Based on the study by Song *et al.* (2016), the ANN approach has been verified as being capable of restoring high temporal resolution rainfall rates based on TBR data after model calibration. Further improvements have been achieved compared with other methods, yet with poor model performance in the validation. This study aims to extend and further explore the ANN model performance in different precipitation groups assigned with varied TBR rainfall intensities at two different sites in the UK. The model was originally built using data from Sparsholt in the UK and then applied to Chilbolton to further test the model extrapolation performance. In addition, a LIM and a CSA are introduced and compared with the ANN method by using the Nash–Sutcliffe efficiency (NSE), and the root mean square error (RMSE) performance indicators.

STUDY AREAS AND DATASET

In this study, the data are collected from Chilbolton (51°08′N, 01°26′W) and Sparsholt (51°04′N, 01°26′W) observatories, which are located approximately 8 km apart in southern England (Figure 1). Regional climatic differences may exist to affect the variability of raindrop size distribution but with slight discrepancies between these two closely located sites (Townsend *et al.* 2009). Therefore, the rainfall distribution has similar patterns and properties in both sites. The disdrometer data are measured by an impact disdrometer of the type JWD RD-69 and the data are available from April 2003 to September 2014 from the Chilbolton site, and from July 2004 to September 2013 from the Sparsholt site. However, during mid-August 2004 to mid-December 2004 and from July 2005 to May 2006, the precipitation instrument in Chilbolton was returned to the manufacturer for maintenance, and no data were available during this period (Islam *et al.* 2012). In contrast, the data from Sparsholt are complete during the whole study period. The data collected during the period 2007 to 2009 have relatively complete measurements for both sites. An RW Munro 0.2 mm TBR was collocated with the

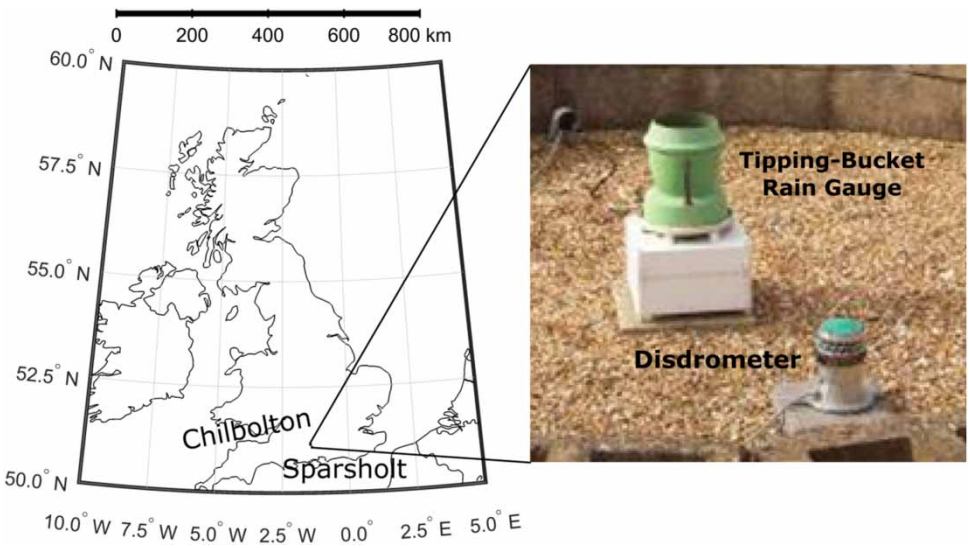

Figure 1 | Locations of the JWD and TBR at Chilbolton and Sparsholt observatories.

Downloaded from http://iwaponline.com/jh/article-pdf/19/6/930/659024/jh0190930.pdf
by guest

disdrometers at both sites and the rainfall measurement period correponds with that of the disdrometers. This dataset is available through the British Atmospheric Data Centre (BADC) website (http://catalogue.ceda.ac.uk/uuid/b5b96f48a8ea9493fedad621dbc1fc5d). Moreover, it is undeniable that the lower the difference between the disdrometer rainfall and the rain gauge measurements, the higher the performance that can be obtained in terms of rainfall rates. The best-correlated data were used because it is important to establish the most reliable models from the best possible data. Poor quality data can result in unreliable models. The minimum annual rainfall amount differences between the TBR and the JWD can be identified in 2009 at Sparsholt and Chilbolton observatories separately (13.4 mm and 4.6 mm, respectively) as shown in Table 1, as large rainfall amount differences, such as those in 2007 (217.4 mm and 158.3 mm) and 2008 (141.2 mm and 111.0 mm), can be prone to measurement errors which introduce uncertainties in rainfall rate comparisons. Consequently, the dataset for 2009 was selected in this study. The detailed information of data pre-processing is presented in the section 'Model setup'.

RESEARCH METHODOLOGY

Disdrometer and TBR measurements

The RD-69 JWD consists of three main components, which are a transducer, a processor and an analogue to digital converter. The original data of JWD is drop size distribution measured every 10 seconds in 127 size bins, which correspond to raindrops of different sizes. The detailed derivation process of rainfall intensity can be found in Joss & Waldvogel (1977); Montopoli *et al.* (2008) and Sheppard (1990). Previous studies have indicated that it is hard to find a universal agreement for the integration time using disdrometer observations. As for a long time period, it may smoothen and miscount the existing physical variations; while for a short time period, the counting fluctuations would dominate the observed raindrop size distribution (Montopoli *et al.* 2008). In this study, the measurements were averaged into 1-minute intervals to filter out time variations, which has been widely used in many previous studies.

The original data of TBRs are tip number measured every 10 seconds and rainfall rate derived from TBRs is to divide the accumulative tip volume amount over the 1-minute time period. It is vital to define what a rainfall event is as the rainfall rates generated from LIM and CSA are notably affected by it. Using CSA based upon the time gap between consecutive tips to define rainfall events is arbitrary to some extent, e.g., 15 minutes (Wang *et al.* 2008) and 10 minutes (Sadler & Busscher 1989). The Tropical Rainfall Measuring Mission Satellite Validation Office (TSVO) defines the end of a rainfall event when there is a time gap larger than 15 minutes (Wang *et al.* 2008). The time scales can significantly affect the rainfall rates generated from TBRs, especially for longer periods, as TBRs are incapable

Table 1 | The annual rainfall amounts in mm between the TBR and the JWD at Sparsholt and Chilbolton observatories from 2007 to 2009; mean and variance of rainfall rates in different groups at Sparsholt and Chilbolton in 2009

	Sparsholt			Chilbolton		
Year	**TBR[a]**	**JWD[a]**	**|JWD-TBR|[a]**	**TBR**	**JWD**	**|JWD-TBR|**
2007	608.8	826.2	217.4	680.8	839.1	158.3
2008	464.0	605.2	141.2	674.6	785.6	111.0
2009	707.0	693.6	13.4	587.8	583.2	4.6
		Mean[a]	**Variance[a]**		**Mean**	**Variance**
2009	Group 1	2.740	3.272	Group 1	1.828	0.778
	Group 2	2.843	6.042	Group 2	2.955	7.969
	Group 3	4.041	18.174	Group 3	3.669	17.993
				Group 4	3.842	29.014

[a]The units of TBR, JWD, and |JWD-TBR| are in mm; mean: mm/hour; variance: $(mm/hour)^2$.

Downloaded from http://iwaponline.com/jh/article-pdf/19/6/930/659024/jh0190930.pdf
by guest

of capturing rainfall variation during low rainfall intensity periods (Habib *et al.* 2001). Based on previous studies and after applying different time scales to define the rainfall events in this study, it was found that by increasing the time of consecutive gaps greater than 15 minutes, the CSA defectively fits the accumulative rainfall amount especially when there is little rainfall during the gap. This is because the CSA requires piecewise continuous property due to its considerable sensitivity to the smoothness of the third or higher derivatives, as mentioned before. Consequently, in this study, the end of a rainfall event is defined when the time gap between consecutive tips is larger than 15 minutes.

ANNs

ANNs are normally depicted as systematic structures of interconnected neurons, each of which has its own activation level and responsibility to propagate the message (i.e., data) from the input layer to the output layer. A typical multi-layer neural network with a feed-forward configuration is shown in Figure 2, where circles represent neurons. Information, along with the weights of each connection, is transmitted and exchanged with each other between the input, hidden and output layers. The weighted inputs with bias from each neuron in the previous layer is transmitted to the neurons in the hidden layer, which can be described as (Tengeleng & Armand 2014):

$$O_j^k = f_k\left(\sum_{i=1}^{N_{k-1}} I_i^{k-1} W_{ij}^k + b_j^k\right) \tag{1}$$

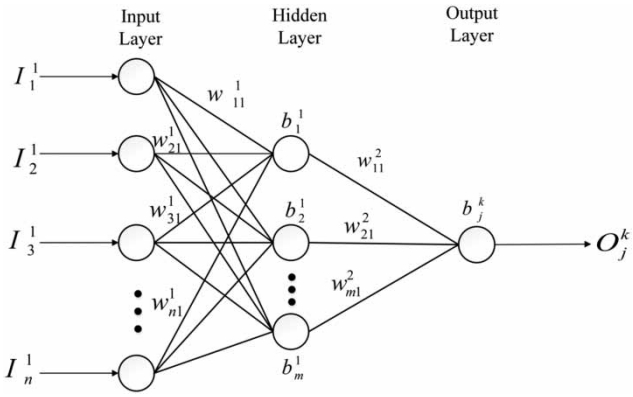

Figure 2 | The schematic representation of ANN.

where I_i^{k-1} is the input from *ith* node in the *(k – 1)th* layer, W_{ij}^{k-1} denotes the weight between node j and all the nodes in the previous layers, b_j^k represents the bias at node j in the *kth* layer, N_{k-1} is the number of nodes in the layer $k-1$, and f_k represents the activation function to model nonlinear behaviours. The weights are adjusted through an iterative training process to minimize the error between theoretical and experimental outputs until they coincide within a given tolerance (Tengeleng & Armand 2014). The typical convergence criteria to assess the performance function is the mean square error, which is given by:

$$C = MSE = \frac{\sum_{i=1}^{N} (T_i - O_i)^2}{N} \tag{2}$$

where T_i and O_i are the target and output values at time i, respectively, N is the number of data points. The detailed description of the backpropagation algorithm is provided in the Appendix (available with the online version of this paper). The Levenberg–Marquardt backpropagation method is used in the training phase since it has the fastest backpropagation and is highly recommended as the first-choice supervised algorithm with the feed-forward neural network (Hagan *et al.* 1996).

The validation/testing process is interpreted when the neural network is deemed to be competent in calculating other responses to new entries that were never used in the training. However, questions may arise on how to partition the data into training and testing datasets. Shahin *et al.* (2004) found that there is no universal rule to separate the data since it depends on data quantity and an optimal model performance can be achieved when 20% of the dataset is used for validation when the sub-datasets have statistical consistent properties. Therefore, this study splits the training and validation datasets into the ratio of 80%–20%, respectively, of the whole data time series after analysing the statistical properties of the datasets to avoid underfitting/overfitting issues.

Model setup

In this study, the ANN model setup is specified as follows: one input layer, one hidden layer and one output layer are configured for the feed-forward ANN. A preliminary analysis confirms that when integrating original data from TBR

Downloaded from http://iwaponline.com/jh/article-pdf/19/6/930/659024/jh0190930.pdf
by guest

as inputs in ANN, the estimated rainfall rates are worse than both LIM and CSA. Although discrepancies of rainfall rates derived from disdrometer, LIM and CSA exist, the overall rainfall rates among them still have strong correlations. It is logical that more precise results can be obtained when input and output data are highly correlated. Consequently, this can be a heuristic idea of using rainfall rates from LIM and CSA as inputs to the ANN in this study. Besides those two types of rainfall rates, the time distance variable, which is defined as the time difference between tips, is also considered as input. This variable is important because it is logical that the shorter the time difference between tips, the heavier the rainfall rate, especially for short-time rainfall events, and vice versa during low-intensity rainfall events. In order to verify how this input impacts on the output, a sensitivity analysis for the number of time distance input variables, as well as their permutations, is applied between inputs and target values. Figure 3 depicts the schematic diagram of the number variation of the time distance input variable. The solid vertical lines represent the tips and the dashed vertical line denotes the reference data point (i.e., the point for the rainfall rate to be estimated). The time distance, i.e., -1 and $+1$, indicates the distance from the reference point to the previous and the following tip, respectively; the others have the same corresponding definitions. However, sensitivity analysis indicates that permutations involved with the previous and the following two tips are more correlated with the target variable. Consequently, six input variables (LIM, CSA, $+1$, -1 $+2$, -2) are used in the ANN in this study, whereas the disdrometer rainfall rate (DIS) is considered as the target variable. As a result, the rainfall rate data are processed as the following steps: first, TBR and DIS are synchronized based on dates which are available for both datasets; then, rainfall events are defined when the time gap between consecutive tips is larger than 15 minutes based on TBR time series; the time distance between consecutive tips is considered as input which further constrains that each predefined rainfall event should contain at least four tips; lastly, those rainfall events are refined by removing the first and last two tips to maintain the same data lengths in all the inputs.

For the decision of hidden layer size, a key issue denoted as the 'bias and variance trade-off' should be specified to obtain the proper number of neurons in the hidden layer. During the training process, the bias and variance are calculated and minimized through each training iteration. However, it is unrealistic to prevent supervised learning algorithms from generalizing beyond the training set by minimizing bias and variance simultaneously. Algorithms with high bias typically produce simpler models; however, they may underfit the training data, thus failing to capture important features in the dataset. On the contrary, high-variance reveals that learning methods may be able to represent the well-trained dataset, but are at risk of overfitting to noise or unrepresentative training data algorithms (Hastie *et al.* 2009). Consequently, it is crucial to find the optimum model complexity where not only the regularities can be accurately recognized in the training data, but also be properly generalized in the test data. To ensure the model has the optimum setup, the *K*-fold cross-validation is used to

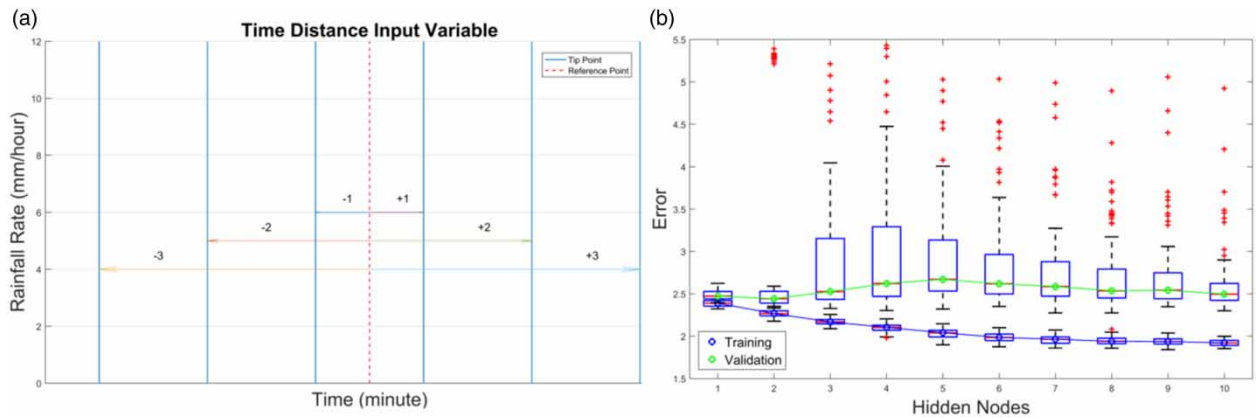

Figure 3 | (a) Schematic diagram of time distance input variable; (b) 100 ensemble five-fold cross-validation tests on the hidden layer size.

Downloaded from http://iwaponline.com/jh/article-pdf/19/6/930/659024/jh0190930.pdf
by guest

comprehensively assess the true accuracy of the system. The basic principle is first to partition the original dataset successively into approximately k equal subsamples. A single subsample is then retained as the validation data for model testing, and the $k-1$ remaining subsamples are used as training data. Finally, the cross-validation process is repeated k times and each of the k subsamples is used once as the validation data (Hastie *et al.* 2009). This study uses five-fold cross-validation based on the data partition principle as mentioned before. Figure 3 shows 100 ensemble five-fold cross-validation analysis on the number of hidden neurons ranging from 1 to 10. The boxplots in both training (lower part) and validation (upper part) datasets depict the error distribution where the lower and upper circles represent the median error values with different hidden neurons. The results reveal that the error decreases during training (lower solid line) with the increase of neurons whereas the error fluctuates (upper solid line) in a wider range during the validation. Moreover, when the model contains more than one neuron in the hidden layer, it is prone to overfitting during the training process and shows poor and unstable performance in the validation data, as shown by the crosses. In addition, the crosses increase notably with the number of hidden layers and this can be due to the issue of 'curse of dimensionality', in which an increase in the inputs will lead to the exponential rise of dimensionality, thus requiring more samples to stop neural networks memorizing the data (Priddy & Keller 2005).

Therefore, one neuron is used in the hidden layer and one correspondent output is achieved in the output layer. After the above data processing steps, the time series lengths for Sparsholt and Chilbolton are 8,385 and 6,454, respectively. When the ANN is trained with more data with a wide range of values, not only is overfitting avoided in the training process, but it also improves the ANN generalization as being able to perform better especially with values within that range (Hagan *et al.* 1996). In this study, the model is calibrated and validated with the Sparsholt dataset and further tested with the Chilbolton dataset to assess its performance.

The agreement of rainfall rates, derived from ANN, CSA and LIM in terms of the reference rainfall rate measured by the disdrometer, is assessed by using the NSE and the RMSE and accumulative rainfall amount (ACCU) as performance indicators. The NSE is dimensionless and the units of RMSE and ACCU are mm/hour and mm, respectively. The theoretical descriptions of NSE and RMSE are described in the Appendix.

It is vital to decide the method of partitioning the dataset in training and validation with respect to the data properties. It is believed that if the ANN model is trained with a wide range of values, the model is expected to perform better with the validation dataset (Shahin *et al.* 2004). Consequently, this study split the whole time series into the training and validation data separately, where the training data cover most of the value spans and the validation data, which are not used during training, are divided into three groups with values ranging from low to high, as shown in Figure 4. Table 1 further illustrates the statistics (mean and variance) of rainfall rates in those three groups. It can be seen that the mean and variance increase from group 1 to group 3, especially for the data in group 3 which are sparsely distributed as shown by the high variance (18.174 (mm/hour)^2), which is about six times greater than the data in group 1 (3.272 (mm/hour)^2). By contrast, group 1 has both the lowest mean (2.740 mm/hour) and variance, which indicates that these data have a narrow distribution with low rainfall intensities.

RESULTS AND DISCUSSION

Assessment of model performance at Sparsholt

Figure 4 shows the input (LIM, CSA) rainfall time series as well as the validation results for 2009 at Sparsholt observatory. The validation results are specified and compared herein to explore the model performance in three groups based on the value range partition.

Group 1

The validation dataset for the first group in Figure 4 has the lowest rainfall rates, which are less than 12 mm/hour, but are widely distributed (mean: 2.740 mm/hour; variance: 3.272 (mm/hour)^2 in Table 1) when compared with other groups. Table 2 lists the model performance results from the training and validation data. It can be concluded from Table 2 that the ANN is much better than LIM and CSA for the training dataset as the NSE and RMSE are 0.858

Downloaded from http://iwaponline.com/jh/article-pdf/19/6/930/659024/jh0190930.pdf
by guest

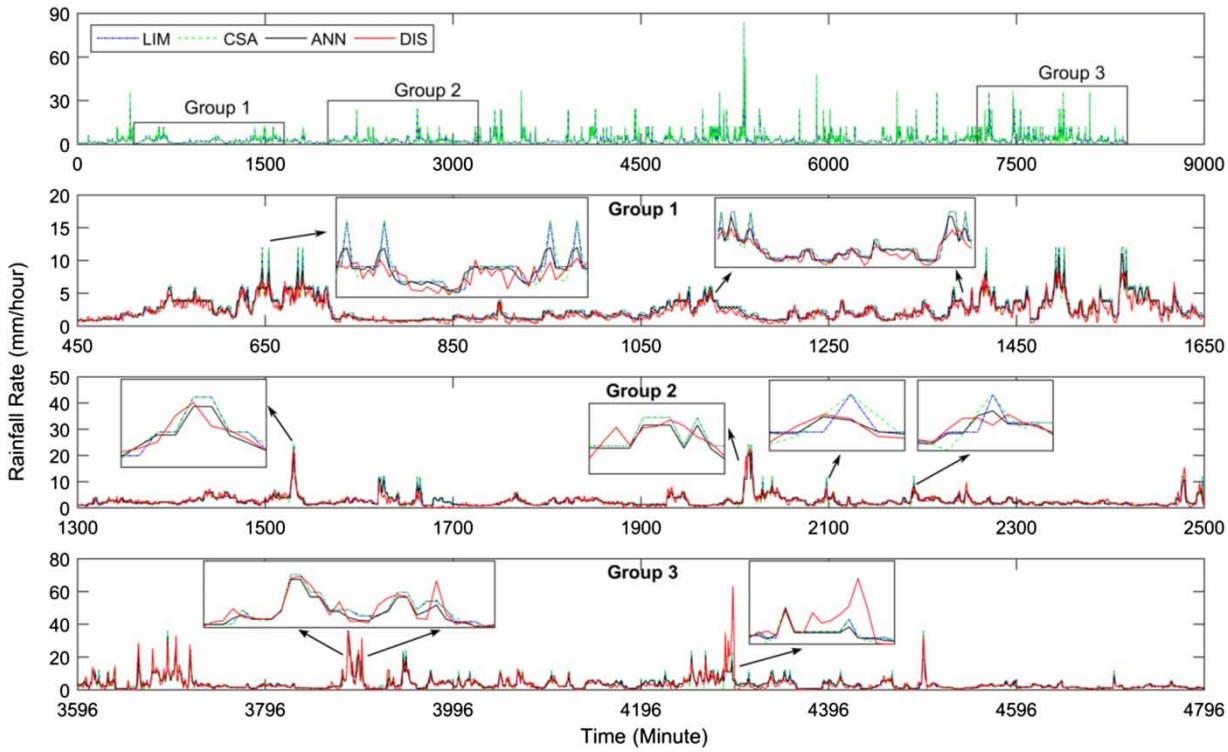

Figure 4 | Time series of input (LIM, CSA) rainfall rates and validation results in 2009 at Sparsholt.

Table 2 | Model performance of 3 groups

				Validation								
	Training			Group 1			Group 2			Group 3		
Methods	NSE	RMSE*	ACCU*	NSE	RMSE	ACCU	NSE	RMSE	ACCU	NSE	RMSE	ACCU
LIM	0.764	1.669	256.700	0.567	0.979	54.790	0.676	1.287	56.827	0.616	2.992	80.796
CSA	0.759	1.689	256.904	0.542	1.006	54.817	0.652	1.334	56.888	0.612	3.005	80.861
ANN	0.858	1.296	230.123	0.757	0.733	49.531	0.762	1.104	50.947	0.663	2.800	72.501
DIS*	1.000	0.000	230.122	1.000	0.000	44.755	1.000	0.000	51.713	1.000	0.000	72.402

*DIS: disdrometer; RMSE: mm/hour; ACCU: mm.

and 1.296 mm/hour separately, which are about 0.10 higher and 0.40 mm/hour lower than those of other groups. Moreover, the ACCU between ANN and DIS is identical, which further indicates the success of ANN in regression analysis during training. Compared with the NSE for the training data of all three methods, the validation NSEs are relatively low (under 0.800). However, the ANN in group 1 distinctly outperforms the other two since the NSE reaches 0.757 and RMSE decays to 0.733 mm/hour (about 0.20 increase in NSE and 0.30 mm/hour decrease in RMSE). Although the ANN slightly overestimates the ACCU (about 5 mm, which is half of the other methods), it still has the closest estimation with the disdrometer (DIS), as listed in Table 2. In addition, the ANN's advantages over LIM and CSA are further illustrated in Figure 4, as it substantially detects the rainfall rate variation of the disdrometer data, while the LIM and CSA cannot, particularly at the points where the input rainfall rate values are equal to 12 mm/hour

Downloaded from http://iwaponline.com/jh/article-pdf/19/6/930/659024/jh0190930.pdf
by guest

around time 650 (zoom-in subfigure) and 1,100 (zoom-in subfigure) in Figure 4.

Group 2

The second group is described as the validation dataset that contains both low and comparatively high rainfall rates with a mean and variance of 2.843 mm/hour and 6.042 (mm/hour)^2, respectively, as shown in Figure 4 and Table 1. Compared with group 1, the ANN consistently performs better than LIM and CSA in terms of higher NSE and lower RMSE although with less improvement (NSE increases from 0.652 to 0.762; RMSE decreases from 1.334 mm/hour to 1.104 mm/hour), as shown in Table 2. However, the ANN constantly outperforms LIM and CSA in the ACCU estimation (50.947 mm) as it has the minimum difference (0.234 mm) with DIS (51.713 mm). The capability of precisely capturing the target rainfall rates for the ANN can be clearly demonstrated in Figure 4, especially at low values points of 2,097 and 2,190 (two right-side zoom-in subfigures) although underestimating rainfall rates at high values at 1,530 and 2,014 (two left-side zoom in subfigures).

Group 3

The validation data with higher rainfall rate values with sparse distribution (mean: 4.041 mm/hour; variance: 18.174 (mm/hour)^2 in Table 1) are defined as group 3 and shown in Figure 4. It can be inferred from Table 2 that the ANN achieves an equivalent performance with LIM and CSA despite a minor increase (0.612 to 0.663) of NSE and decrease (3.005 mm/hour to 2.800 mm/hour) of RMSE, whereas it has the best estimation of ACCU because of the small difference with DIS. The validation results of the ANN in Figure 4 indicates its ability to trace the target rainfall rates' values ranging from low to intermediate rainfall rates, such as at points 3,850 (left-side zoom-in subfigure) where rainfall rates are lower than 40 mm/hour. Nevertheless, the ANN along with LIM and CSA fails to estimate the peak rainfall rate at point 4,296 (right-side zoom-in subfigure) in Figure 4, where the maximum target rainfall rate occurs and this is probably due to the low correlation between the input and target values, which leads to poor estimation with the ANN.

From the results of the model performance at Sparsholt, the ANN has more advantages in retrieving high temporal rainfall rates than the other methods. Moreover, since rainfall has similar patterns in both observatories, the models built at Sparsholt are supposed to provide accurate rainfall rates as well at Chilbolton. Thus, another test was carried out with the Chilbolton observatory dataset by directly applying the models developed for Sparsholt, to further verify the feasibility of the ANN's advantages. Figure 5 depicts the time series of input (LIM, CSA) rainfall rates and the validation results for the year 2009 at Chilbolton observatory. It is worth specifying that a fourth group (group 4) was introduced in the test dataset for Chilbolton as it covers most of the largest values that cannot be applied in the Sparsholt dataset because of the partition principle of training and validation. Based on the principle of data selection at Sparsholt, Table 1 illustrates the statistics (mean and variance) of rainfall rates in four groups at Chilbolton. Similarly, the mean and the variance increase from group 1 to group 4, however, the variance between each group jumps dramatically from 0.778 (mm/hour)^2 (group 1) to 29.014 (mm/hour)^2 (group 4) which indicates data (excludes group 1) considerably scattered among the groups though with small difference of mean values.

Assessment of model performance at Chilbolton

Group 1

Rainfall rate data in group 1, overall, have low magnitude and a wide distribution of rainfall values with a mean and variance of 1.828 mm/hour and 0.778 (mm/hour)^2, respectively, as shown in Table 1. Comparing Table 2 with Table 3 for group 1, the results reveal that the ANN provides better rainfall rate predictions than LIM and CSA. For instance, the NSE increased from 0.595 to 0.646 and the RMSE decreased from 0.750 mm/hour to 0.701 mm/hour, which indicates a slight improvement with the ANN. For the ACCU estimation, the ANN performs better due to the lowest bias when compared with DIS. Figure 5 manifests the advantages of the ANN approach as it roughly estimates the high rainfall values, as shown in the time steps from 650 to 780 (left-side zoom-in subfigure), although overestimates the disdrometer rainfall intensities from time 1,100 to

Downloaded from http://iwaponline.com/jh/article-pdf/19/6/930/659024/jh0190930.pdf
by guest

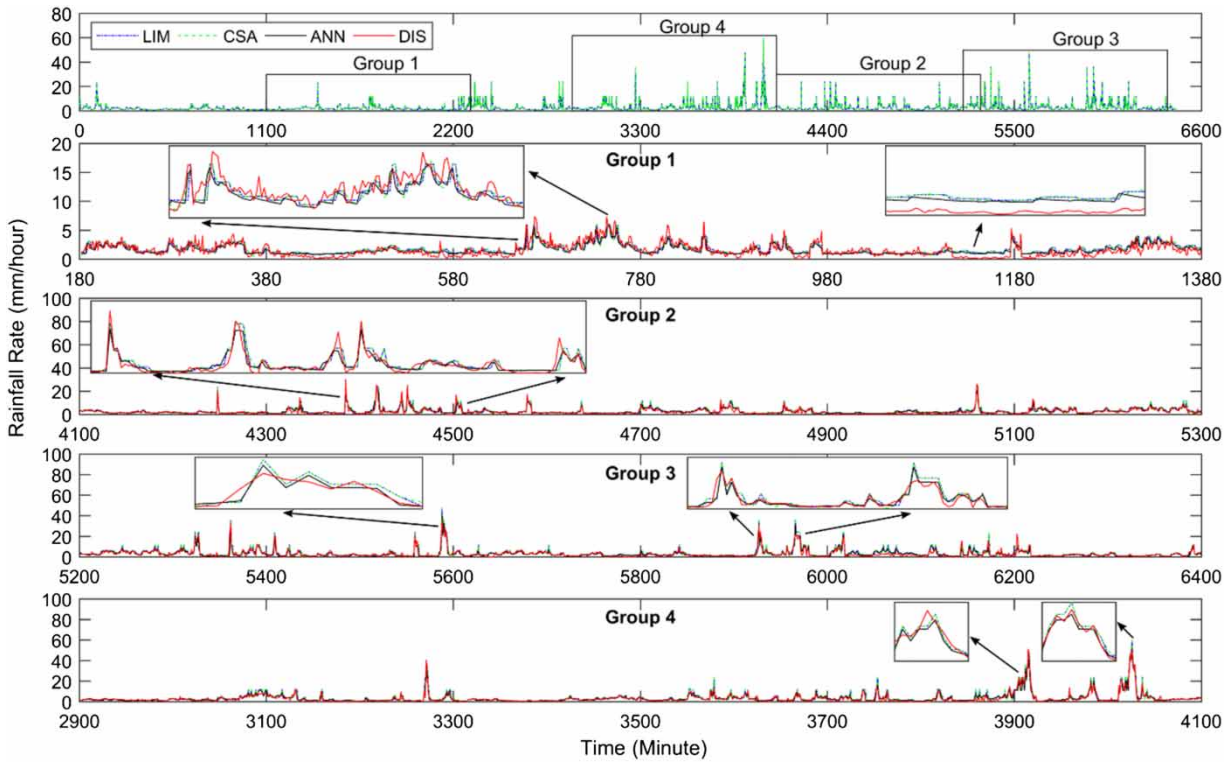

Figure 5 | Time series of input (LIM, CSA) rainfall rates and validation results for 2009 at Chilbolton.

Table 3 | Model performance of four groups

	Test											
	Group 1			Group 2			Group 3			Group 4		
Methods	NSE	RMSE*	ACCU*	NSE	RMSE	ACCU	NSE	RMSE	ACCU	NSE	RMSE	ACCU
LIM	0.595	0.750	36.549	0.701	1.540	59.021	0.640	2.325	73.350	0.809	2.200	76.796
CSA	0.609	0.737	36.561	0.702	1.537	59.173	0.641	2.324	73.416	0.811	2.190	76.877
ANN	0.646	0.701	32.566	0.788	1.298	53.387	0.772	1.850	65.536	0.890	1.672	68.087
DIS*	1.000	0.000	32.404	1.000	0.000	47.892	1.000	0.000	61.084	1.000	0.000	64.866

*DIS: disdrometer; RMSE: mm/hour; ACCU: mm.

1,170 (right-side zoom-in subfigure). This also illustrates the shortcomings of the ANN as the model outputs are predominantly dependent on the input data quality.

Group 2

The rainfall rates selected in group 2 show lower rainfall values (mean: 2.955 mm/hour) and relatively scattered distribution (variance: 7.969 $(mm/hour)^2$), as shown in Table 1. The ANN is better than LIM and CSA (high NSE: 0.788; low RMSE: 1.298 mm/hour; minimum ACCU difference: 5.495 mm), as listed in Table 3 for group 2. Moreover, the ANN presents its strengths in capturing variations of rainfall rates from low to high values at time steps 4,380–4,510 (zoom-in subfigure), as shown in Figure 5.

Downloaded from http://iwaponline.com/jh/article-pdf/19/6/930/659024/jh0190930.pdf
by guest

Group 3

Data selected in group 3 incorporate more high values with a higher mean (3.669 mm/hour) and variance (17.993 (mm/hour)^2), as shown in Table 1. The results shown in Table 3 for group 3 imply that the ANN noticeably exceeds the other two groups with higher NSE (increases from 0.640 to 0.772), lower RMSE (decreases from 2.325 mm/hour to 1.850 mm/hour) and smallest ACCU differences (4.452 mm). Figure 5 enhances the evidence that the ANN is able to predict the disdrometer rainfall rates well when its values are lower than 30 mm/hour, albeit it shows overestimation for higher values at time step 5,590 (left-side zoom-in subfigure), 5,950 and 5,980 (right-side zoom-in subfigure).

Group 4

Group 4, overall, contains the highest rainfall rates (mean: 3.842 mm/hour) and most dispersedly distributed (variance: 29.014 (mm/hour)^2) rainfall rate data, as shown in Table 1. It can be concluded from group 4 that the NSE for all methods are all above 0.80 and the ANN reached 0.889 along with the closest ACCU (68.087 mm) compared with DIS (64.866 mm), as shown in Table 3. The superior estimation from the ANN is further confirmed in Figure 5 as most of the target rainfall values have a good performance with the ANN rainfall predictions, especially at higher values as shown in time step 3,915 (left-side zoom-in subfigure) and time step 4,030 (right-side zoom-in subfigure).

Based on this analysis of test results in four groups in Chilbolton, it can be summarized that the ANN is able to inherently provide the best estimation with LIM and CSA methods, especially when input and target rainfall rates are highly correlated, as described in group 4. In addition, the ANN exhibits its intrinsic strengths as well as its consistency of stability and capability in accurate high temporal rainfall rate estimation in both sites.

CONCLUSIONS

This study compared three methods: ANNs, CSA and a LIM to retrieve rainfall rates at high temporal resolutions. Rainfall rates are derived from TBRs located at Chilbolton and Sparsholt observatories in the UK for the year 2009. The ANN model was integrated with LIM and CSA rainfall rates as well as the time difference between the tips as inputs to the ANN model. The NSE and RMSE were used for the assessment of the model performance.

The model results have shown that the ANN substantially surpasses CSA and LIM with the validation datasets integrated with different input rainfall rate values ranging from low to high based upon the model trained within a wide span of rainfall values. Furthermore, the ANN model has these advantages along with consistent and stable abilities to retrieve precise high temporal rainfall rates at nearby sites. Consequently, it is recommended that such a model should be built at small spatial scales in the current phase by using disdrometers and other rainfall measurement instruments that are able of capturing rainfall intensities with high temporal resolution. Additionally, the ANN model high temporal resolution rainfall rates, corresponding with those derived from disdrometer and TBRs, can be helpful in remote sensing research fields in the calibration and validation of radar and satellite-borne precipitation measurements at short temporal scales. However, the limitations of this study also present the poor performance of rainfall estimation by the ANN approach when the input variables have low correlation with the target rainfall variable as well as the partition of the training and testing datasets based on data quality and quantity. Moreover, it should be emphasized that more rainfall data and rainfall events in other places with different climatic conditions should be explored using the proposed method to build a more robust neural network model for its wide generalization not only locally but also distantly. Eventually, since rainfall is a stochastic variable that changes in space and time, this study herein provides an idea to explore and evaluate the rainfall rate agreement derived from the JW disdrometer and TBRs with an ANN approach and conventional methods, which can be useful in many meteorological and hydrological applications.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the British Atmospheric Data Centre (BADC) for providing the dataset used in this study (http://catalogue.ceda.ac.uk/uuid/b5b96f48a8ea9493fedad621dbc1fc5d). We also thank

Downloaded from http://iwaponline.com/jh/article-pdf/19/6/930/659024/jh0190930.pdf
by guest

the anonymous reviewers for providing insightful comments that helped to improve the manuscript.

REFERENCES

Bringi, V. N. & Chandrasekar, V. 2001 The polarimetric basis for characterizing precipitation. In: *Polarimetric Doppler Weather Radar*. Cambridge University Press, Cambridge, pp. 378–533. doi:10.1017/CBO9780511541094.009.

Garcia-Pintado, J., Barberá, G. G., Erena, M. & Castillo, V. M. 2009 Rainfall estimation by rain gauge-radar combination: a concurrent multiplicative-additive approach. *Water Resour. Res.* **45**. doi:10.1029/2008WR007011.

Habib, E., Krajewski, W. F. & Kruger, A. 2001 Sampling errors of tipping-bucket rain gauge measurements. *J. Hydrol. Eng.* **6**, 159–166. doi:10.1061/(ASCE)1084-0699(2001)6:2(159).

Hagan, M. T., Demuth, H. B. & Beale, M. H. 1996 *Neural Network Design*. PWS Pub, Boston, MA, p. 734.

Hastie, T., Tibshirani, R. & Friedman, J. 2009 The elements of statistical learning. *Elements* **1**, 337–387. doi:10.1007/b94608.

Islam, T., Rico-Ramirez, M. A., Han, D. & Srivastava, P. K. 2012 A Joss-Waldvogel disdrometer derived rainfall estimation study by collocated tipping bucket and rapid response rain gauges. *Atmos. Sci. Lett.* **13**, 139–150. doi:10.1002/asl.376.

Jaffrain, J. & Berne, A. 2011 Experimental quantification of the sampling uncertainty associated with measurements from PARSIVEL disdrometers. *J. Hydrometeorol.* **12**, 352–370.

Joss, J. & Waldvogel, A. 1977 Comments on 'Some observations on the Joss-Waldvogel rainfall disdrometer'. *J. Appl. Meteorol.* **16**, 112–113. doi:10.1175/1520-0450(1977)016<0112:COOOTJ>2.0.CO;2.

Kirstetter, P.-E., Gourley, J. J., Hong, Y., Zhang, J., Moazamigoodarzi, S., Langston, C. & Arthur, A. 2015 Probabilistic precipitation rate estimates with ground-based radar networks. *Water Resour. Res.* **51**, 1422–1442. doi:10.1002/2014WR015672.

Liu, X. C., Gao, T. C. & Liu, L. 2013 A comparison of rainfall measurements from multiple instruments. *Atmos. Meas. Tech.* **6**, 1585–1595. doi:10.5194/amt-6-1585-2013.

Montopoli, M., Marzano, F. S. & Vulpiani, G. 2008 Analysis and synthesis of raindrop size distribution time series from disdrometer data. *Geosci. Remote Sensing, IEEE Trans.* **46**, 466–478. doi:10.1109/TGRS.2007.909102.

Ochoa-Rodriguez, S., Wang, L. P., Gires, A., Pina, R. D., Reinoso-Rondinel, R., Bruni, G., Ichiba, A., Gaitan, S., Cristiano, E., Van Assel, J., Kroll, S., Murlà-Tuyls, D., Tisserand, B., Schertzer, D., Tchiguirinskaia, I., Onof, C., Willems, P. & Ten Veldhuis, M. C. 2015 Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: a multi-catchment investigation. *J. Hydrol.* **531**, 389–407. doi:10.1016/j.jhydrol.2015.05.035.

Priddy, K. L. & Keller, P. E. 2005 Artificial neural networks: an introduction. SPIE Press.

Rico-Ramirez, M. A., Bringi, V. N. & Thurai, M. 2012 River flow simulations with polarimetric weather radar. Moore, R. J., Cole, S. J. & Illingworth, A. J. (eds). In: Weather Radar and Hydrology (Proceedings of a symposium held in Exeter, UK, April 2011). vol. 351, IAHS Press, Exeter, pp. 466–471.

Sadler, E. J. & Busscher, W. J. 1989 High-intensity rainfall rate determination from tipping-bucket rain gauge data. *Agron. J.* doi:10.2134/agronj1989.00021962008100060016x.

Savina, M., Schäppi, B., Molnar, P., Burlando, P. & Sevruk, B. 2012 Comparison of a tipping-bucket and electronic weighing precipitation gage for snowfall. *Atmos. Res.* **103**, 45–51. doi:10.1016/j.atmosres.2011.06.010.

Shahin, M. A., Maier, H. R. & Jaksa, M. B. 2004 Data division for developing neural networks applied to geotechnical engineering. *J. Comput. Civ. Eng.* **18**, 105–114. doi:10.1061/(ASCE)0887-3801(2004)18:2(105).

Sheppard, B. E. 1990 Effect of irregularities in the diameter classification of raindrops by the Joss-Waldvogel disdrometer. *J. Atmos. Ocean. Technol.* **7**, 180–183. doi:10.1175/1520-0426(1990)007<0180:EOIITD>2.0.CO;2.

Sideris, I. V., Gabella, M., Erdin, R. & Germann, U. 2014 Real-time radar-rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland. *Q. J. R. Meteorol. Soc.* **140**, 1097–1111. doi:10.1002/qj.2188.

Sieck, L. C., Burges, S. J. & Steiner, M. 2007 Challenges in obtaining reliable measurements of point rainfall. *Water Resour. Res.* **43**. doi:10.1029/2005WR004519.

Song, Y., Han, D. & Rico-Ramirez, M. A. 2016 High temporal resolution rainfall information retrieval from tipping-bucket rain gauge measurements. *Procedia Eng.* **154**, 1193–1200. doi:10.1016/j.proeng.2016.07.525.

Tengeleng, S. & Armand, N. 2014 Performance of using cascade forward back propagation neural networks for estimating rain parameters with rain drop size distribution. *Atmosphere (Basel)* **5**, 454. doi:10.3390/atmos5020454.

Thurai, M. & Bringi, V. N. 2005 Drop axis ratios from a 2D video disdrometer. *J. Atmos. Ocean. Technol.* **22**, 966–978.

Tokay, A., Bashor, P. G. & Wolff, K. R. 2005 Error characteristics of rainfall measurements by collocated Joss-Waldvogel disdrometers. *J. Atmos. Ocean. Technol.* **22**, 513–527. doi:10.1175/JTECH1734.1.

Townsend, A. J., Watson, R. J. & Hodges, D. D. 2009 Analysis of the variability in the raindrop size distribution and its effect on attenuation at 20–40 GHz. *IEEE Antennas Wirel. Propag. Lett.* **8**, 1210–1213. doi:10.1109/LAWP.2009.2035724.

Wang, J., Fisher, B. L. & Wolff, D. B. 2008 Estimating rain rates from tipping-bucket rain gauge measurements. *J. Atmos. Ocean. Technol.* **25**, 43–56. doi:10.1175/2007JTECHA895.1.

First received 9 April 2017; accepted in revised form 22 June 2017. Available online 24 August 2017

Downloaded from http://iwaponline.com/jh/article-pdf/19/6/930/659024/jh0190930.pdf
by guest