Forecasting Hourly Intermittent Rainfall by Deep Belief Networks with Simple Exponential Smoothing

Guo-Yu Huang · Chi-Ju Lai · Ping-Feng Pai

Abstract

Accurate rainfall forecasting is essential in planning and managing water resource systems efficiently. However, intermittent rainfall patterns increase the difficulty of accurately forecasting rainfall values. Deep learning techniques have recently been popular and powerful in forecasting. Thus, this study employed deep belief networks with a simple exponential smoothing procedure (DBNSES) to forecast hourly intermittent rainfall values in Taiwan. Weather factors were used as independent variables to forecast rainfall volume. The simple exponential smoothing data preprocessing procedure was used to deal with the intermittent data patterns. The other three forecasting models, namely the least squares support vector regression (LSSVR), the generalized regression neural network (GRNN), and the backpropagation neural network (BPNN), were employed to forecast rainfall using the same data sets. In addition, genetic algorithms were utilized to determine the parameters of four forecasting models. The empirical results indicate that the developed DBNSES models are superior to the other forecasting models in terms of forecasting accuracy. In addition, the DBNSES can obtain smaller values of RMSE than those in the previous studies. Therefore, the DBNSES model is a suitable and effective way of forecasting rainfall with intermittent data patterns.

Keywords     Intermittent rainfall · Deep belief networks · Forecast

1 Introduction

In many countries, accurately forecasting intermittent rainfall is a major challenge, particularly in Taiwan, whose topographical elements and geographic position pose difficulty (Wu and Lin 2017). Due to the mountainous terrains and the island nature in Taiwan, most rainfall occurs in the plum rain season, including May and June, and the typhoon seasons, from July to October. Heavy downpours frequently result in catastrophic repercussions like...
flooding, landslides, floods, or debris flows that cause property damage and fatalities (Saito et al. 2017; Song and Park 2020). Nevertheless, extensive studies have gained valuable insight into building reliable warning and forecasting systems to alleviate disasters threatening security and the economy because of excessive rainfall. Through proper preventive measures, such as evacuating people living in the most endangered areas, provided by early warning systems, the influence of disasters resulting from heavy rainfall could be alleviated. Thus, accurate rainfall forecasting plays a crucial role in disaster prevention and mitigation in Taiwan.

In hydrological research, precisely predicting rainfall is a crucial topic, particularly in countries with intermittent weather patterns (Wu and Lin 2017; Singh 2018; Haidar and Verma 2018) because accurate rainfall forecasts can moderate the impacts of heavy rain and prevent disasters timely. The forecasts of hourly rainfall in hydrology are acquired from radar measurement or by using various statistical methods to forecast rainfall, including machine learning, singular spectrum analysis, support vector machine, and time series. Yu et al. (2015) developed a hybrid blending system that collects and integrates information from radar data and weather data to improve the forecasting accuracy of floods and rainfall. It was reported that the designed model could generate more accurate and effective forecasting results than the other radar-based forecasting methods. Yu et al. (2017) employed random forest techniques and support vector machine models with radar-derived data to forecast real-time rainfall. Previous grid positions, including longitude and latitude and three reservoirs, Feitsui, Deji, and Zengwen, served as independent variables to forecast rainfall 1 to 3 h ahead. This investigation reported that the support vector machine approach is superior to the random forest in terms of forecasting accuracy. Bagirov et al. (2017) used clusterwise linear regression to forecast rainfall. Data employed in this study were gathered from eight atmospheric stations. Five weather attributes—maximum temperature, minimum temperature, evaporation, vapor pressure, and solar radiation—were treated as independent variables of the clusterwise linear regression to forecast monthly rainfall. This investigation reported that the clusterwise linear regression outperformed the other four prediction models. Unnikrishnan and Joshiprakash (2018a) performed a singular spectrum analysis to forecast daily rainfall in time series. Various components of the rainfall series, such as trend, periodic component, noise, and cyclic component, were extracted. Numerical results indicated that the presented model could effectively forecast the daily rainfall in time series for long durations. Further, Ghamariadyan and Imteaz (2021) developed a wavelet artificial neural networks model to forecast seasonal rainfall in Queensland, Australia. This study reported that the presented model is superior to the other five forecasting models by more than 35% according to forecasting accuracy. Islam and Imteaz (2022) designed a hybrid autoregressive integrated moving average with an exogenous input model to predict seasonal rainfall in Western Australia. Numerical results revealed that the proposed model outperformed the other traditional linear and non-linear models in terms of forecasting accuracy.

Deep learning is an emerging technology applied in many fields, including forecasting rainfall. Venkata Ramana et al. (2013) employed the artificial neural network with the wavelet technique to forecast monthly rainfall in Darjeeling. Three variables were selected for independent variables, namely rainfall, minimum temperature, and maximum temperature. Numerical results revealed that the proposed wavelet neural network outperformed the artificial neural network. Azad et al. (2015) investigated the predictability of monsoon rainfall in India using monthly and annual data. In addition, wavelets and artificial neural networks...
were applied to extract the periodic structure and train forecasting models, respectively. The results indicated that more accurate results are provided when periodic and random components are considered separately. Salman et al. (2015) compared the performances of three deep learning methods, namely recurrent neural networks, conditional restricted Boltzmann machines, and convolutional neural networks, in rainfall forecasting. Numerical results revealed that the recurrent neural networks could obtain more accurate rainfall forecasting results than the other two methods. Shi et al. (2015) developed a convolutional long short-term memory network model to nowcast rainfalls. The designed method can effectively forecast rainfall intensity over a relatively short period of time by spatiotemporal data. Hernández et al. (2016) designed a deep learning architecture composed of two networks—the autoencoder network and the multilayer perceptron network—to forecast the accumulated rainfall for the following day. The results indicated that the proposed forecasting model outperformed multilayer perceptron neural networks. Ha et al. (2016) used deep belief networks to model with eight-fold cross-validation in their study. Data were collected from an automatic weather station, including rainfall, temperature, and sun and moon positions. The numerical results showed that the proposed model outperformed the multilayer perceptron neural networks in terms of forecasting accuracy. Dabral and Murry (2017) employed Seasonal Autoregressive Integrative Moving Average models (SARIMA) on three types of data sets to forecast the monsoon rainfall, including monthly, weekly, and daily time series data. The autocorrelation function, partial autocorrelation function, and the minimum values of the Akaike Information Criterion with Schwarz Bayesian Information were used to select the most optimal SARIMA models. Further, Qiu et al. (2017) designed a multi-task convolutional neural network to forecast short-time rainfall. Due to the correlations between different sites that can effectively forecast rainfall in various weather conditions, this investigation used weather datasets from multiple stations for training. The proposed models were significantly superior to European Centre for Medium-Range Weather Forecasts. Liu et al. (2022) proposed the convolutional long short-term memory network model to nowcast the rainfall of four stations in Helan Mountain, namely Dianjiang, Shihuiyao, Linkuang, and Suyukou, using rainfall data and radar data. Combined reflectance and the retrieved wind field served as the main determinants for forecasting. Compared with convolutional neural networks and long short-term memory networks, the proposed model greatly improved the performance of forecasting rainfall with the ability to capture spatial information and timing memory. Wei and You (2022) presented a hybrid model by integrating the discrete wavelet transform with two deep learning models, namely dilated causal convolutional neural network and long short-term memory, to forecast the monthly rainfall of four cities in China, specifically Beijing, Tianjin, Chongqing, and Guangzhou. The experimental results indicated that the proposed model outperformed long short-term memory and dilated the causal convolutional neural network. Further, it was confirmed to be more satisfactory for long-term rainfall forecasting.

The rest of this paper is organized as follows: Sect. 2 introduces deep belief networks and the framework of the DBNSES model in rainfall forecasting; Sect. 3 presents the numerical results; Sect. 4 addresses the conclusions and direction of the future study.
2 The Proposed DBNSES Model in Rainfall Forecasting

2.1 Deep Belief Networks

A deep belief network (DBN), proposed by Hinton et al. (2006a), includes several stacked restricted Boltzmann machines (RBMs)—the main components of a deep belief network—and a back propagation neural network (BPNN). The RBM was a revised form of the original Boltzmann machine (BM) (Ackley et al. 1985) with node connections both within layers and between layers. The difference lies in that the RBM only allows links between nodes in adjacent layers. The learning strategies of a DBN can be split into two phases, including unsupervised learning in the pre-training phase and supervised learning in the fine-tuning phase. In the pre-training phase, the unsupervised learning–based training uses only the independent variable to select apposite initial parameters, namely weights and biases, for the supervised learning. Therefore, to maximize the likelihood estimation, the pre-training phase rebuilds training samples by adjusting parameters. According to the initial parameters determined by the pre-training phase, supervised learning in the fine-tuning phase is performed to further efficiently adjust the network parameters. The outputs of the units in a visible layer served as inputs of the units in the hidden layer. Equation (1) illustrates the energy function of a connected structure \((\varrho, \varsigma)\) of visible and hidden units.

\[
E_{\text{ngy}}(\varrho, \varsigma) = - \sum_{i} \alpha_{i}\varrho_{i} - \sum_{j} \beta_{j}\varsigma_{j} - \sum_{i} \sum_{j} \varrho_{i}w_{ij}\varsigma_{j}
\]  

where \(W_{ij}\) is the weight connecting unit \(i\) in the visible layer with a bias \(\alpha_{i}\) and unit \(j\) in the hidden layer with a bias \(\beta_{j}\). Values of \(\varrho_{i}\) and \(\varsigma_{j}\) are binary distributions and can be depicted as Eq. (2):

\[
p(\varrho, \varsigma) = \frac{1}{z} e^{\text{ngy}(\varrho, \varsigma)}
\]

where \(z = \sum_{\varrho, \varsigma} e^{\text{ngy}(\varrho, \varsigma)}\) is a normalization partition function. Activated probability functions are expressed as Eq. (3) and Eq. (4), while the state vector \(\varrho\) of the visible layer or \(\varsigma\) of the hidden layer is produced.

\[
p(\varrho_{i} = 1|\varsigma) = \text{sigmoid} \left( \sum_{j} w_{ij}\varsigma_{j} + \alpha_{i} \right)
\]

\[
p(\varsigma_{j} = 1|\varrho) = \text{sigmoid} \left( \sum_{i} w_{ij}\varrho_{i} + \beta_{j} \right)
\]

The \(\text{sigmoid}(•)\) is a sigmoid function represented as Eq. (5):

\[
\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}}
\]
Three parameters, $\alpha$, $\beta$, and $W$, are illustrated as a parameter set $\theta$. Basically, the objective of RBM is to pursue an appropriate parameter set. The outcomes of unsupervised RBM are treated as input of the MLP, which is the supervised learning phase of deep belief networks. When distributions of the visible and the hidden layers are in a steady phase, the likelihood function is maximized and the features of the training data set can be represented well using a constructive divergence algorithm (Hinton 2002b; Roux and Bengio 2008).

Equation (6) depicts the likelihood function:

$$\ln L(\theta) = \sum_{s=1}^{s} \ln P(\varrho_s)$$  \hfill (6)$$

where $s$ is the size of training data. The gradient of Eq. (6) is illustrated as Eq. (7) in regard to the parameter set.

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{s=1}^{s} \frac{\partial \ln P(\varrho_s)}{\partial \theta}$$  \hfill (7)$$

Equation (8) and Eq. (9) are the results of the derivative of $\ln P(\varrho_s)$ in regard to $\theta$ and the gradient of the likelihood function, correspondingly.

$$\frac{\partial \ln P(\varrho_s)}{\partial \theta} = -\sum_{s} P(\varsigma|\varrho_{s}) \frac{\partial \text{Energy}(\varrho_s, \varsigma)}{\partial \theta} + \sum_{s,s} P(\varrho, \varsigma) \frac{\partial \text{Energy}(\varrho, \varsigma)}{\partial \theta}$$  \hfill (8)$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{s=1}^{s} \left(-\sum_{s} P(\varsigma|\varrho_{s}) \frac{\partial \text{Energy}(\varrho_s, \varsigma)}{\partial \theta} + \sum_{s,s} P(\varrho, \varsigma) \frac{\partial \text{Energy}(\varrho, \varsigma)}{\partial \theta} \right)$$  \hfill (9)$$

Thus, the learning rate and the momentum of the RBM can be represented as Eq. (10).

$$\theta(t) = m_p \theta(t-1) + \lambda_p \frac{\partial \ln L(\theta)}{\partial \theta}$$  \hfill (10)$$

where $\lambda_p$ and $m_p$ denote the learning rate and the momentum, correspondingly. Then, the k-step constructive divergence algorithm (Hinton 2002b) was used to cope with the computations of Eq. (10). In this study, the k value was equal to one. The training policy of parameters in RBM was depicted as Eqs. (11)–(13).

$$w_{ij}^{t+1} = m_p^* w_{ij}^t + \Delta w_{ij}, \Delta w_{ij} = \lambda_p \left(P(\varsigma|\varrho^{(0)}) [\varrho^{(0)}]^T - P(\varsigma|\varrho^{(1)}) [\varrho^{(1)}]^T \right)$$  \hfill (11)$$

$$\alpha_{ij}^{t+1} = m_p^* \alpha_{ij}^t + \Delta \alpha_{ij}, \Delta \alpha_{ij} = \lambda_p (\varrho^{(0)} - \varrho^{(1)})$$  \hfill (12)$$

$$\beta_{ij}^{t+1} = m_p^* \beta_{ij}^t + \Delta \beta_{ij}, \Delta \beta_{ij} = \lambda_p \left(P(\varsigma|\varrho^{(0)}) - P(\varsigma|\varrho^{(1)}) \right)$$  \hfill (13)$$

The well-trained parameters provided by RBM served as initial parameters of the next supervised learning stage of deep belief networks. Backpropagation algorithms were used for training weights and biases of multilayer perceptron networks in the supervised learning stage. The output of the hidden layer can be expressed as Eq. (14):
$$s^l_j = f \left( \sum_{i=1}^{n} w^l_{ij} x_i - b^l_j \right), \ j = 1, 2, \ldots, z$$  \hspace{1cm} (14)$$

where \( n \) and \( z \) are the numbers of neurons in the input layer and hidden layer, respectively. The number of the hidden layer is \( l \). The \( w^l_{ij} \) is the weight between the input layer and the hidden layer. \( x_i \) is the input variable, namely the weather factors, and \( b \) is the bias. The output of the output layer was presented as Eq. (15).

$$Y_k = f \left( \sum_{j=1}^{l} s^l_{jk} w^l_{jk} - b^l_k \right), \ k = 1, 2, \ldots, m$$  \hspace{1cm} (15)$$

where \( Y_k \) is the forecasting value, namely rainfall, in this study. The mean square error, illustrated as Eq. (16), was used to measure the error between actual values and forecasting values.

$$E_k = \frac{1}{2} \sum_{k} (A_k - Y_k)^2$$  \hspace{1cm} (16)$$

where \( A_k \) is the actual value. Integrating Eqs. (14) and (15) into Eq. (16), and Eq. (17) was obtained.

$$E_k = \frac{1}{2} \sum_{k} \left( A_k - f \left( \sum_{j=1}^{l} f \left( \sum_{i=1}^{n} w^l_{ij} x_i - b^l_j \right) w^l_{jk} - b^l_k \right) \right)^2$$  \hspace{1cm} (17)$$

The errors between the hidden layer and the output layer were calculated by the delta function expressed as Eq. (18).

$$\delta^l_k = \begin{cases} 
(A_k - Y_k) f'(\sum_{j=1}^{l} O_j w^l_{jk} - b^l_k), & l = l + 2 \\
(\sum_{i=1}^{l-1} \delta^{l-1}_{ij} w^{l-1}_{ij}) f'(\sum_{j=1}^{l} O_j w^l_{jk} - b^l_k), & 2 \leq l \leq l + 2 
\end{cases}$$  \hspace{1cm} (18)$$

Therefore, the learning rules of the weights and the biases in the supervised learning phase were presented as Eqs. (19)–(21):

$$w^{l+1}_{jk} = w^l_{jk} + \Delta w_{jk} = w^l_{jk} + \eta \delta_k O_j$$  \hspace{1cm} (19)$$

$$w^{l+1}_{ij} = w^l_{ij} + \Delta w_{ij} = w^l_{ij} + \eta \delta_k x_i$$  \hspace{1cm} (20)$$

$$b^{l+1}_k = b^l_k + \eta \delta_k$$  \hspace{1cm} (21)$$

where \( \eta \) is the learning rate of the backpropagation (Huang 2019).
2.2 The Framework of DBNSES Model in Rainfall Forecasting

In this study, a deep belief network with a simple exponential smoothing model was designed to forecast hourly intermittent rainfall. Three other forecasting models—the least squares support vector regression, generalized regression neural networks, and the backpropagation neural networks—were also employed to forecast rainfall with the same data sets to demonstrate the performance of the presented DBNSES model. The flowchart of this study is illustrated in Fig. 1.

The intermittent rainfall frequently leads to floods and inundations, which result in property damage and personal injury. New Taipei City is Taiwan’s most populated county, with Banqiao District having the highest population. To ensure public safety and lessen the likelihood of flooding, drainage systems and early warning systems were improved. In addition, effective water resource management systems and accurate rainfall forecasting are considered essential in preventing drought and flood disasters. First, the weather attributes and hourly rainfall data at Banqiao District, New Taipei City were collected from the Central Weather Bureau in Taiwan (https://e-service.cwb.gov.tw/HistoryDataQuery). Figure 2 illustrates the geographic region of this study. In total, ten weather attributes were used, including station pressures, sea pressures, temperatures, dew-point temperatures, relative humidity, wind speeds, wind directions, max gusts, the directions of max gusts, and rainfall hours. Rainfalls in Taiwan have an intermittent characteristic. Thus, simple exponential smoothing was used to preprocess the intermittent rainfall data. Previous studies (Waller 2015; Kourentzes 2013; Li and Lim 2018) pointed out that the simple exponential smoothing represented by Eq. (22) could smoothen intermittent data. Hence, the simple exponential smoothing method served as a preprocessing procedure in this study:

\[
F_t = F_{t-1} + \alpha (R_{t-1} - F_{t-1})
\]

where \(\alpha\) is a smoothing constant with values between 0 and 1; \(R_t\) and \(F_t\) represent the actual value and the forecasting value at time \(t\), respectively. The weather attributes and rainfall data served as independent variables and the dependent variable, correspondingly. After applying the simple exponential smoothing method to the rainfall data, ten weather attributes and rainfall were normalized before being fed into the deep belief networks. The data were divided into training data and testing data for model learning and forecasting performance evaluations. Both the 24 h-ahead and the 168 h-ahead predictions were conducted in this study to examine the performance of the proposed DBNSES model in forecasting hourly intermittent rainfall. The numbers of the training data set and testing data set of the 24 h-ahead forecastings are 8736 and 8784, correspondingly. In addition, the amount of the training data set and testing data set of the 168 h-ahead forecastings are 8592 and 8784, respectively.

Furthermore, this study used genetic algorithms (Holland 1975) to determine the parameters of four forecasting models—DBNSES, LSSVR, GRNN, and BPNN. Based on the study of Lin et al. (2019), the proposed DBNSES model employed genetic algorithms to select six parameters, including the learning rate and the momentum of both unsupervised and supervised training phases, and the dropout and the number of neurons in the hidden layer in the supervised learning phase. Table 1 illustrates parameters and searching ranges for forecast-
Fig. 1 The flowchart of the DBNSES model for hourly intermittent rainfall forecasting
ing models. For unsupervised learning and supervised learning phases, the energy function of restricted Boltzmann machines and root mean square error values of rainfall forecast were treated as the fitness functions of genetic algorithms, respectively. Both minimizations of the energy function and the root mean square error were pursued in the learning process of deep belief networks. Thus, negative outputs of both measurements were employed as the fitness function of genetic algorithms for maximizations. The gene number of a chromosome, the number of generations, the population size, the crossover, and the mutation rates of genetic algorithms are 40, 30, 30, 0.7, and 0.7, respectively. Further, this study applied a single-point crossover with binary coding for genetic algorithms. When the search for appropriate parameters of DBNSES models was completed, the finalized DBNSES models were obtained. The finalized forecasting models were then used to predict rainfall. Then, the forecasting rainfall values were restored by the reversed functions of the normalization and the simple exponential smoothing. Finally, measurements of rainfall forecasting accuracy were generated. Algorithm 1 shows the process of the proposed DBNSES models with the

Table 1 Parameters and searching ranges of forecasting models

| Models  | Parameters                                           | Searching ranges    |
|---------|------------------------------------------------------|---------------------|
| DBNSES  | Learning rates of RBM and MLP                        | [0.1, 0.9]          |
|         | Momentum of RBM and MLP                              | [0.1, 0.9]          |
|         | Dropout of MLP                                       | [0.1, 0.9]          |
|         | Number of neurons in the hidden layer                 | [2, 50]             |
| LSSSVR  | Penalty coefficients and kernel function parameters  | [1, 500]            |
| GRNN    | Spread parameters                                    | [0.1, 0.9]          |
| BPNN    | Learning rates and momentum                           | [0.1, 0.9]          |

Fig. 2 The geographic region of the study
Genetic Algorithm. In this study, the proposed forecasting models were implemented by the MATLAB and the spreadsheet. The MATLAB were used to implement forecasting models. The data were collected and arranged in spreadsheets. The designed forecasting system is able to read weather and rainfall data then make a forecast.

Algorithm 1: The proposed DBNSES with genetic algorithms

1. **Input**: dataset with 10 weather attributes and rainfall;
2. Preprocess rainfall data by simple exponential smoothing;
3. Normalize weather attributes and rainfall;
4. Divide data into training datasets and testing datasets;
5. Initialize a DBNSES parameter set with six parameters in a form of population: $\text{POP}_i$, ($i = 1, 2, \ldots, n$);
6. While end condition ($i = n$) of genetic algorithms is not reached do
   <The unsupervised training phase>
7. Repeat until the number of epochs is reached for training RBM;
8. Train parameter set $\{\alpha, \beta, W\}$ by k-step constructive divergence algorithm;
9. Use the output of the layer RBM parameter set as the input of next layer RBM;
10. Stop;
11. Use well-trained weights provided by RBM as initial weights of the BPNN;
<The supervised training phase>
12. Repeat until the number of epochs is reached for training BPNN;
13. Use gradient descent method to train weights;
14. Stop;
15. Calculate $\text{RMSE}_i$;
16. Conduct genetic algorithms and generate a new set of six parameters of DBNSES: $\text{POP}_{i+1}$;
17. if ($\text{RMSE}_{i+1} < \text{RMSE}_i$) then
18. Replace $\text{POP}_i$ with $\text{POP}_{i+1}$
19. else
20. remain the $\text{POP}_i$
21. end if;
22. end while;
23. Generate finalized DBNSES models;
24. Calculate the forecasting accuracy by testing data.

3 Numerical Results

Four forecasting models—deep belief networks, the least squares support vector regression, generalized regression neural networks, and backpropagation neural networks—were employed to forecast intermittent rainfall with the same data sets. Tables 2 and 3 list the parameters of the four models for 24 h-ahead and 168 h-ahead forecastings. In this study, the range of smoothing constants ranged from 0.1 to 0.9 interval value of 0.1. The smoothing constants resulting in the smallest RMSE were selected. For the 24 h-ahead forecastings, the smoothing constants of simple exponential smoothing leading to the smallest RMSE values for DBN, LSSVR, GRNN, and BPNN models were 0.1, 0.9, 0.9, and 0.1, correspondingly. For the 168 h-ahead forecastings, the smoothing constants of simple exponential smoothing leading to the smallest RMSE values for DBN, LSSVR, GRNN, and BPNN models were 0.9, 0.9, 0.9, and 0.8, respectively. Two indices, the root mean square error (RMSE) and the mean absolute error (MAE), expressed as Eqs. (23) and (24), were used to measure the performance of the forecasting models:
\[
\text{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^{N} (F_t - A_t)^2}
\]  
\[
\text{MAE} = \frac{1}{N} \sum_{t=1}^{N} |F_t - A_t|
\]  

where \( N \) is the number of forecasting periods, \( F_t \) is the forecasting value at time \( t \), and \( A_t \) is the actual value at time \( t \). Table 4 lists the forecasting measurements with the best \( \alpha \) by four models for two forecasting strategies. The forecasting measurements without simple exponential smoothing by four models for two forecasting strategies are shown in Table 5. Tables 4 and 5 show that the hourly intermittent rainfall forecasting performance can be mostly improved using a simple exponential smoothing preprocessing approach, particularly in 24 h-ahead forecastings. The DBNSES models were superior to the other three models of both 24 h-ahead and 168 h-ahead forecastings in terms of RMSE. It was pointed out that RMSE is more sensitive to outliers than MAE (Hyndman and Koehler 2006). In addition, the proposed DBNSES model outperformed previous studies of rainfall forecasting in terms of RMSE (Wu and Lin 2017; Singh 2018; Unnikrishnan and Jothiprakash 2018b). Thus,
the numerical results revealed that the proposed DBNSES is a feasible and stable method for forecasting hourly intermittent rainfall. Additionally, the simple exponential smoothing technique can yield satisfactory results when coping with data with intermittent patterns. Figures 3 and 4 plot the absolute error values of rainfall provided by forecasting models for two different forecasting strategies, respectively. The proposed DBNSES can deliver more satisfactory results than the other three forecasting models. Figure 5 illustrates the absolute error values of rainfall generated by forecasting models with and without simple exponential smoothing in a format of boxplots. A boxplot can demonstrate error values and provide clear data distributions when a large number of values in the visual representation arises (Tukey 1977). It can be observed from Fig. 5 that the simple exponential smoothing technique reduced the median values and outliers in most forecasting models, especially for the deep belief networks.

4 Conclusion

Although rainfall resources are quite abundant in Taiwan, its economic progress is hampered by frequent and heavy rains. In recent years, extreme weather has had a devastating impact on people and caused a variety of calamities globally. Due to its topographical features and geographic position, Taiwan has a convergent and intermittently heavy rainfall characteristic. Thus, effective rainfall forecasting can help the public and commercial sectors take the right precautions to lessen the risk of disasters. Therefore, this study developed a deep belief network with a simple exponential smoothing model for forecasting hourly rainfall in Taiwan. Both rainfall data and weather factors were used to forecast rainfall. Numerical results indicate that DBNSES models could outperform the other three models in both 24 h-ahead and 168 h-ahead forecastings in terms of RMSE. In addition, the proposed DBNSES model can generate smaller values of RMSE than those in previous studies in rainfall forecasting. Compared to models with and without simple exponential smoothing, it was shown that the simple exponential smoothing technique could deliver more reliable results and positive influence when dealing with intermittent data. Thus, the proposed DBNSES models are a feasible and promising alternative to hourly rainfall forecasting. Noteworthily, this study

| Forecasting strategies | Measurements | Forecasting models |
|------------------------|--------------|--------------------|
|                        | DBN | LSSVR | GRNN | BPNN |
| 24 h ahead             | RMSE | 1.642 | 1.666 | 2.410 | 2.559 |
|                        | MAE  | 0.450 | 0.445 | 0.446 | 1.077 |
| 168 h ahead            | RMSE | 1.642 | 1.668 | 2.351 | 9.364 |
|                        | MAE  | 0.446 | 0.452 | 0.503 | 4.829 |

Table 4 The measurements of forecasting models with the simple exponential smoothing generating the smallest RMSE value for each model

| Forecasting strategies | Measurements | Forecasting models |
|------------------------|--------------|--------------------|
|                        | DBN | LSSVR | GRNN | BPNN |
| 24 h ahead             | RMSE | 22.376 | 1.796 | 2.720 | 27543.340 |
|                        | MAE  | 8.702 | 0.481 | 0.469 | 1129.978 |
| 168 h ahead            | RMSE | 1.65 | 2.022 | 2.272 | 1.803 |
|                        | MAE  | 0.229 | 0.601 | 0.505 | 0.633 |

Table 5 The forecasting measurements without the simple exponential smoothing by four models for two forecasting strategies
only used weather attributes for rainfall forecasting. For future study, some other factors, such as radar maps (Kim et al. 2017), could be included in the DBNSES to increase the accuracy of rainfall forecasting. In the narrowest sense, weather and rainfall data collected from a weather station were used to examine the effectiveness of the proposed DBNSES models. In the broadest sense, weather and rainfall data gathered from other regions or countries can be used to investigate the feasibility of the designed DBNSES models.

Furthermore, the findings of this study are as follows. First, weather attributes are essential in forecasting intermittent rainfall when deep belief networks are used. Second, simple exponential smoothing is an effective way to improve forecasting results for deep belief networks, and the genetic algorithm is a feasible and promising way to determine the parameters of deep belief networks. Finally, with the advantages of simple exponential smoothing preprocessing, genetic algorithms, and deep learning capabilities, the developed DBNSES models can be appropriately used for contexts with intermittent rainfall.

Fig. 3 Absolute errors of forecasting models with the smoothing constant generating the smallest RMSE for 24 h ahead forecasting
Fig. 4 Absolute errors of forecasting models with the smoothing constant generating the smallest RMSE for 168 h ahead forecasting.
Fig. 5 Absolute error values of rainfall in a format of boxplots

Authors' Contributions G.-Y. Huang: Data collection, data analysis, model development; C.-J. Lai: Draft paper writing, visualization; P.-F. Pai: Conceptualization, model development, final paper writing, funding acquisition.
Funding This research was partially funded by the Ministry of Science and Technology, Taiwan, under the Contract Number MOST 109-2410-H-260-023.

Data Availability Data used in this study will be made available upon request.

Data Availability Available on request.

Declarations

Ethical Approval Not applicable.

Consent to Participate All authors have consented.

Consent to Publish All authors have consented.

Competing Interests All authors declare that they have no conflict of interest.

References

Ackley DH, Hinton GE, Sejnowski TJ (1985) A learning algorithm for Boltzmann machines. Cognit Sci 9:147–169

Azad S, Debnath S, Rajeevan M (2015) Analysing predictability in Indian monsoon rainfall: a data analytic approach. Environ Processes 2:717–727

Bagirov AM, Mahmood A, Barton A (2017) Prediction of monthly rainfall in Victoria, Australia: Clusterwise linear regression approach. Atmos Res 188:20–29

Dabral PP, Murry MZ (2017) Modelling and forecasting of rainfall time series using SARIMA. Environ Processes 4:399–419

Ghamariyad M, Imteaz MA (2021) Prediction of Seasonal Rainfall with One-year Lead Time Using Climate Indices: A Wavelet Neural Network Scheme. Water Resour Manage 35:5347–5365

Haidar A, Verma B (2018) Monthly rainfall forecasting using one-dimensional deep convolutional neural network. IEEE Access 6:69053–69063

Hernández E, Sanchez-Anguix V, Julian V, Palanca J, Duque N (2016) Rainfall prediction: A deep learning approach. International Conference on Hybrid Artificial Intelligence Systems 151–162

Ha JH, Lee YH, Kim YH (2016) Forecasting the rainfall of the next day using deep learning. J Korean Inst Intell Syst 26:93–98

Hinton GE, Osindero S, Teh YW (2006a) A fast learning algorithm for deep belief nets. Neural Comput 18:1527–1800

Hinton GE (2002b) Training products of experts by minimizing contrastive divergence. Neural Comput 14:1711–1800

Huang GY (2019) A study on deep belief networks in rainfall forecasting. Unpublished Master Dissertation National Chi Nan University

Holland JH (1975) Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. MIT Press, pp 439–444

Hyndman RJ, Koehler AB (2006) Another look at measures of forecast accuracy. Int J Forecast 22:679–688

Islam F, Imteaz MA (2022) A Novel Hybrid Approach for Predicting Western Australia’s Seasonal Rainfall Variability. Water Resour Manage

Kim S, Hong S, Joh M, Song SK (2017) DeepRain: Convlstm network for rainfall prediction using multi-channel radar data. arXiv preprint arXiv:1711.02316.

Kourentzes N (2013) Intermittent demand forecasts with neural networks. Int J Prod Econ 143:198–206

Li C, Lim A (2018) A greedy aggregation–decomposition method for intermittent demand forecasting in fashion retailing. Eur J Oper Res 269:860–869

Lin KP, Pai PF, Ting YJ (2019) Deep belief networks with genetic algorithms in forecasting wind speed. IEEE Access 7:99244–99253

Liu W, Wang Y, Zhong D, Xie S, Xu J (2022) ConvLSTM Network-Based Rainfall Nowcasting Method with Combined Reflectance and Radar-Retrieved Wind Field as Inputs. Atmosphere 13:411
Qiu M, Zhao P, Zhang K, Huang J, Shi X, Wang X, Chu W (2017) A short-term rainfall prediction model using multi-task convolutional neural networks. 2017 IEEE international conference on data mining (ICDM) 395–404

Roux NL, Bengio Y (2008) Representational power of restricted Boltzmann machines and deep belief networks. Neural Comput 20:1631–1649

Saito H, Honda H, Kawahara R (2017) Disaster avoidance control against heavy rainfall. IEEE INFOCOM 2017-IEEE Conference on Computer Communications 1–9

Salman AG, Kanigoro B, Heryadi Y (2015) Weather forecasting using deep learning techniques. 2015 international conference on advanced computer science and information systems (ICACSIS) 281–285

Shi X, Chen Z, Wang H, Yeung DY, Wong WK, Woo WC (2015) Convolutional LSTM network: A machine learning approach for rainfall nowcasting. Advances in neural information processing systems 28

Singh P (2018) Indian summer monsoon rainfall (ISMR) forecasting using time series data: a fuzzy-entropy-neuro based expert system. Geosci Front 9:1243–1257

Song Y, Park M (2020) A study on setting disaster-prevention rainfall by rainfall duration in urban areas considering natural disaster damage. Focusing on South Korea Water 12:642

Tukey JW (1977) Exploratory data analysis. 2:131–160

Unnikrishnan P, Jothiprakash V (2018a) Daily rainfall forecasting for one year in a single run using Singular Spectrum Analysis. J Hydrol 561:609–621

Unnikrishnan P, Jothiprakash V (2018b) Data-driven multi-time-step ahead daily rainfall forecasting using singular spectrum analysis-based data pre-processing. J Hydroinformatics 20:645–667

Venkata Ramana R, Krishna B, Kumar SR, Pandey NG (2013) Monthly rainfall prediction using wavelet neural network analysis. Water Resour Manage 27:3697–3711

Waller D (2015) Methods for Intermittent Demand Forecasting

Wei M, You XY (2022) Monthly rainfall forecasting by a hybrid neural network of discrete wavelet transformation and deep learning. Water Resources Management 1–16

Wu MC, Lin GF (2017) The very short-term rainfall forecasting for a mountainous watershed by means of an ensemble numerical weather prediction system in Taiwan. J Hydrol 546:60–70

Yu W, Nakakita E, Kim S, Yamaguchi K (2015) Improvement of rainfall and flood forecasts by blending ensemble NWP rainfall with radar prediction considering orographic rainfall. J Hydrol 531:494–507

Yu PS, Yang TC, Chen SY, Kuo CM, Tseng HW (2017) Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting. J Hydrol 552:92–104

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.