Rainfall time series forecasting based on Modular RBF Neural Network model coupled with SSA and PLS

Jiansheng Wu Yu
School of Information Engineering, Wuhan University of Technology, P. R. China
Department of Mathematics and Computer, Liuzhou Teacher College, P. R. China
wjsh2002168@163.com

Jimin Yu
School of Automation Institute, ChongQing University of Posts and Telecommunications, P. R. China
Key Laboratory Network Control and Intelligent Instrument, ChongQing University of Posts and Telecommunications, P. R. China
zdyjm@163.com

Abstract:
Accurate forecast of rainfall has been one of the most important issues in hydrological research. Due to rainfall forecasting involves a rather complex nonlinear data pattern; there are lots of novel forecasting approaches to improve the forecasting accuracy. In this paper, a new approach using the Modular Radial Basis Function Neural Network (M–RBF–NN) technique is presented to improve rainfall forecasting performance coupled with appropriate data–preprocessing techniques by Singular Spectrum Analysis (SSA) and Partial Least Square (PLS) regression. In the process of modular modeling, SSA is applied for the time series extraction of complex trends and structure finding. In the second stage, the data set is divided into different training sets by Bagging and Boosting technology. In the third stage, the modular RBF–NN predictors are produced by a different kernel function. In the fourth stage, PLS technology is used to choose the appropriate number of neural network ensemble members. In the final stage, least squares support vector regression is used for ensemble of the M–RBF–NN to prediction purpose. The developed RBF–NN model is being applied for real time rainfall forecasting and flood management in Liuzhou, Guangxi. Aimed at providing forecasts in a near real time schedule, different network types were tested with the same input information. Additionally, forecasts by M–RBF–NN model were compared to the convenient approach. Results show that the predictions made using the M–RBF–NN approach are consistently better than those obtained using the other method presented in this study in terms of the same measurements. Sensitivity analysis indicated that the proposed M-RBF-NN technique provides a promising alternative to rainfall prediction.

Keywords: Singular Spectrum Analysis, Radial Basis Function Neural Network, Partial Least Square Regression, Rainfall prediction, Least Squares Support Vector Regression

1. Introduction
Accurate and timely rainfall prediction is essential for planning and management of water resources, in particular for flood warning systems because it can provide information which help prevent casualties and damage caused by natural disasters [1]. For example, a flood warning system for fast responding catchments may require a quantitative rainfall forecast to increase the lead time for warning. Similarly, a rainfall forecast provide information in advance for many water quality problems [2]. Rainfall prediction is one of the most complex
elements of the hydrology cycle and at the same time difficult to understand and to model due
to the complexity of the atmospheric processes involved and the variability of rainfall in space
and time [3], [4].

Although a physically based approach for rainfall forecasting has had several advantages
in recent decades, given the short time scale, the small catchment area, and the massive costs
associated with collecting required meteorological data, it is not a feasible alternative in most
cases. Over the past few decades, many studies have been conducted for the quantitative
rainfall forecasting using empirical models including multiple linear regression [5], time series
methods [6] and K–nearest–neighbor [7], and data–driven models including artificial neural
network (ANN) [8], support vectors regression (SVR) [9] and fuzzy inference system [10].

Recently, the concept of coupling different models has been a very popular research topic
in hydrologic forecasting, which has attracted scientists from other fields including Statistics,
Machine Learning and so on. They can be broadly categorized into ensemble models and
modular (or hybrid) models. The basic idea behind ensemble models is to build several differ-
et or similar models for the same process and to integrate them together. Their success largely
arises from the fact that they lead to an improved accuracy compared to a single classification
or regression model. Typically, ensemble methods comprise two phases: a) the production
of multiple predictive models, and b) their combination. In recent work, the reduction of the
ensemble size has been the main point of concern [11] [12].

In this paper, unlike the previous work, one of the main purposes is to develop a Modular
Radial Basis Function Neural Network (MRBF–NN) coupled with appropriate data–prepro-
cessing techniques by Singular Spectrum Analysis (SSA) and Partial Least Squares (PLS) to
improve the accuracy of rainfall forecasting. The rainfall data of Liuzhou in Guangxi is pre-
dicted as a case study for our proposed method. An actual case of forecasting monthly rainfall
is illustrated to show the improvement in predictive accuracy and capability of generalization
achieved by our proposed MRBF–NN model.

The rest of this study is organized as follows. Section 2 describes the proposed MRBF–NN,
ideas and procedures. For further illustration, this work employs the method to set up a pre-
diction model for rainfall forecasting in Section 3. Discussions are presented in Section 4 and
conclusions are drawn in the final Section.

2. The building process of the Modular Radial Basis Function Neural
Network

Firstly, Singular Spectrum Analysis (SSA) is used to reduce noises in original rainfall
time series, and to reconstruct the new time series in this section. Secondly, a triple–phase
nonlinear modular RBF–NN model is proposed for rainfall forecasting based on different
activation function and training data. Then an appropriate number of RBF–NN predictors are
selected from the considerable number of candidate predictors by the Partial Least Square
technology. Finally, selected RBF–NN predictors are combined into an aggregated neural
predictor in terms of LS–SVR.

2.1. Singular Spectrum Analysis

The Singular Spectrum Analysis (SSA) technique is a novel and powerful technique of
time series analysis incorporating the elements of classical time series analysis, multivariate
statistics, multivariate geometry, dynamical systems and signal processing method. Broom-
head and King [13] was presented SSA because they show that the singular value decomposi-
tion (SVD) is effective in reducing noises. The aim of SSA is to make a decomposition of the original series into the sum of a small number of independent and interpretable components such as a slowly varying trend, oscillatory components and a structure with less noise [14].

The basic SSA algorithm has two stages: decomposition and reconstruction. The decomposition stage requires embedding and singular value decomposition (SVD). Embedding decomposes an original time series into the trajectory matrix; SVD turns a trajectory matrix into the decomposed trajectory matrices which will turn into the trend, seasonal, monthly components, and white noises according to their singular values. The reconstruction stage demands the grouping to make subgroups of the decomposed trajectory matrices and diagonal averaging to reconstruct the new time series from the subgroups. The SSA algorithm is described in more detail by the related literature [15] [16].

2.2. Radial Basis Function Neural Network

Radial basis function was introduced into the neural network literature by Broomhead and Lowe [17] [18], which was motivated by the presence of many local response neurons in human brain. On the contrary to the other type of NN used for nonlinear regression, like back propagation feed forward networks, it learns quickly and has a more compact topology. The architecture is presented in Figure 1.

The network is generally composed of three layers: an input layer, a single layer of nonlinear processing neuron and output layer. The output of the RBF–NN is calculated according to

$$y_i = f_i(x) = \sum_{k=1}^{N} w_{ik} \phi_k(\|x - c_k\|_2), i = 1, 2, \cdots, m$$

(1)

where $x \in \mathbb{R}^{n \times 1}$ is an input vector, $\phi_k(\cdot)$ is a function from $\mathbb{R}^+$ to $\mathbb{R}$, $\| \cdot \|_2$ denotes the Euclidean norm, $w_{ik}$ are the weights in the output layer, $n$ is the number of neurons in the hidden layer, and $c_k \in \mathbb{R}^{n \times 1}$ are the centers in the input vector space. The functional form of $\phi_k(\cdot)$ is assumed to have been given, and some typical choices are shown in Table 1.

The training procedure of the RBF networks is a complex process. This procedure requires the training of all parameters including the centers of the hidden layer units ($c_i, i = 1, 2, \cdots, m$), the widths ($\sigma_i$) of the corresponding Gaussian functions, and the weights ($\omega_i, i = 0, 1, \cdots, m$) between the hidden layer and output layer. In this paper, the orthogonal least squares algorithm (OLS) is used to train RBF based on the minimizing of SSE. More detailed about the algorithm are provided by the related literature [19].
### Table 1. Types of kernel function name and formula

<table>
<thead>
<tr>
<th>Modual</th>
<th>Functional name</th>
<th>Function formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Linear function</td>
<td>( \phi(x) = x )</td>
</tr>
<tr>
<td>B</td>
<td>Cubic approximation</td>
<td>( \phi(x) = x^3 )</td>
</tr>
<tr>
<td>C</td>
<td>Thin-plate-spline function</td>
<td>( \phi(x) = x^2 \ln x )</td>
</tr>
<tr>
<td>D</td>
<td>Guassian function</td>
<td>( \phi(x) = \exp(-x^2/\sigma^2) )</td>
</tr>
<tr>
<td>E</td>
<td>Multi-quadratic function</td>
<td>( \phi(x) = \sqrt{x^2 + \sigma^2} )</td>
</tr>
<tr>
<td>F</td>
<td>Inverse multi-quadratic function</td>
<td>( \phi(x) = \frac{1}{\sqrt{x^2 + \sigma^2}} )</td>
</tr>
</tbody>
</table>

#### 2.3. Selecting appropriate ensemble members

When data has completed the training, each Modular RBF-NN predictor has generated its own result. However, if there is a great number of individual members, we need to select a subset of representatives in order to improve ensemble efficiency. In this paper, the Partial Least Square (PLS) regression technique is adopted to select appropriate ensemble members.

Partial least squares (PLS) regression analysis was developed in the late seventies by Herman O. A. Wold [20]. PLS regression is particularly useful when we need to predict a set of dependent variables from a (very) large set of independent variables (i.e., predictors). Interested readers can be referred to [21] for more details.

#### 2.4. Least Squares Support Vector Regression

Support vector regression (SVR) was derived form support vector machine (SVM) technique. LS–SVM is a least squares modification to the Support Vector Machine [22]. When SVM can be used for spectral regression purpose, it is called least squares support vector regression (LS–SVR). The major advantage of LS–SVM is that it is computationally very cheap while it still possesses some important properties of the SVM. One of the advantages of LS–SVR is its ability to model nonlinear relationships. In this section we will briefly discuss the LS–SVR method for a regression task. For more detailed information see [23].

Where \( \{x_i, i = 1, 2, \ldots, N\} \) are the output of linear and nonlinear forecasting predictors, \( \{y_i, i = 1, 2, \ldots, N\} \) are the aggregated output and the goal is to estimate a regression function \( f \). Basically we define a \( N \) dimensional function space by defining the mappings \( \varphi = [\varphi_1, \varphi_2, \ldots, \varphi_N]^T \) according to the measured points. The LS-SVM model is of the form \( f(x) = \omega^T \varphi(x) + b \) where \( \omega \) is a weight vector and \( b \) is a bias term. The optimization problem is the following:

\[
\begin{align*}
\min J(\omega, \epsilon) & = \frac{1}{2} \omega^T \omega + \gamma \frac{1}{2} \sum_{i=1}^{N} \epsilon_i^2 \\
\text{s.t. } y_i & = \omega^T \varphi(x_i) + b + \epsilon_i, i = 1, 2, \ldots, N
\end{align*}
\]

where the fitting error is denoted by \( \epsilon_i \). Hyper-parameter \( \gamma \) controls the trade-off between the smoothness of the function and the accuracy of the fitting. This optimization problem leads to a solution by solving the linear Karush–Kuhn–Tucker (KKT) [24]:

\[
\begin{bmatrix}
0 & I_n^T \\
I_n & K + \gamma^{-1}I
\end{bmatrix}
\begin{bmatrix}
b_0 \\
b
\end{bmatrix}
=
\begin{bmatrix}
0 \\
y
\end{bmatrix}
\]
where $I_n$ is a $[n \times 1]$ vector of ones, $T$ means transpose of a matrix or vector, $\gamma$ a weight vector, $b$ regression vector and $b_0$ is the model offset. $K$ is kernel function. A common choice for the kernel function is the Gaussian function:

$$K(x, x_i) = e^{-\frac{|x-x_i|^2}{2\sigma^2}}$$

(4)

2.5. The establishment of Modular RBF–NN

To summarize, the proposed Modular RBF–NN model consists of five main steps. In the process of modular modeling, firstly, SSA is applied for the time series extraction of complex trends and structure finding. Secondly, the data set is divided into different training sets by using Bagging and Boosting technology. Thirdly, the modular RBF–NN predictors are produced by a different kernel function. Fourthly, PLS technology is used to choose the appropriate number of neural network ensemble members. Finally, LS–SVR is used for ensemble of the M-RBF-NN to predict purpose. The basic flowchart diagram can be shown in Figure 2.

Figure 2. The Flowchart of the Modular RBF–NN

3. Results and discussion

3.1. Empirical data

Liuzhou is one of the highly developing cities in southwest of China, and the capital and commercial city of Guangxi. Historical monthly rainfall data was collected from 24 stations of the Liuzhou Meteorology Administration (LMA) rain gauge networks for the period from 1949 to 2010. After analyzing data, the period from January 1949 to December 2006 was selected to train MRBF-NN models, and the data from January 2007 to December 2010 were used as a testing set. Thus the training data set contained 696 data points in time series for MRBF–NN learning, and the other 48 data were used to test sample for MRBF–NN Generalization ability. Fig.3 shows the average monthly rainfall, taken over a period from 1949 to 2010, in Liuzhou. There is one peak of rainfall during a year, in August.
3.2. Criteria for evaluating model performance

Three different types of standard statistical performance evaluation criteria were employed to evaluate the performance of various models developed in this paper. These are average absolute relative error (AARE), root mean square error (RMSE), and the Pearson Relative Coefficient (PRC) which are found in many paper [7].

According to the aforementioned literature, there is a variety of methods for rainfall forecasting model in the past studies. The author used Eviews statistical packages to formulate the ARIMA model. Akaike information criterion (AIC) was used to determine the best model. The model is generated from the data set is AR(5). The equation used is presented in Equation 5.

\[ x_t = 1 - 0.30x_{t-1} - 0.02x_{t-2} - 0.11x_{t-3} + 0.91x_{t-4} + 0.05x_{t-5} \]  

For the purpose of comparison by the same four input variables, we have also built other three rainfall forecasting models: multi-layer perceptron neural network (MLP–NN) model, single RBF–NN and Stacked Regression (SR) ensemble [25] method based on RBF–NN.

The standard RBF–NN were trained for each training set with Gaussian-type activation functions in hidden layer, then tested as an ensemble for each method for the testing set. Each network was trained using the neural network toolbox provided by Matlab software package. In addition, the best single RBF neural network using cross-validation method [21] (i.e., select the individual RBF network by minimizing the RMSE on cross-validation) is chosen as a benchmark model for comparison.

3.3. Analysis of the results

Table 2 illustrates the fitting and testing accuracy and efficiency of the model in terms of various evaluation indices for 696 training and 48 testing samples. From the Table 2, we can generally see that learning ability of M–RBF–NN outperforms the other four models under the same network input. As a consequence, poor performance indices in terms of AARE, RMSE and PRC of AR(5) model is the worst in four model. Table 2 also shows that the performance of M–RBF–NN is the best in case study for training samples.

The more important factor to measure performance of a method is to check its forecasting ability of testing samples in order to apply it to an actual rainfall forecasting. Table 2 shows
Table 2. Performance statistics of the five models for rainfall fitting and forecasting.

<table>
<thead>
<tr>
<th>Model</th>
<th>AR(5)</th>
<th>MLP-NN</th>
<th>RBF-NN</th>
<th>SR</th>
<th>M-RBF-NN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Training data (from 1949 to 2006)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AARE</td>
<td>92.90</td>
<td>82.63</td>
<td>83.84</td>
<td>61.74</td>
<td>52.64</td>
</tr>
<tr>
<td>RMSE</td>
<td>87.85</td>
<td>72.14</td>
<td>73.14</td>
<td>56.69</td>
<td>44.25</td>
</tr>
<tr>
<td>PRC</td>
<td>0.8403</td>
<td>0.8939</td>
<td>0.8901</td>
<td>0.9239</td>
<td>0.9612</td>
</tr>
<tr>
<td><strong>Testing data (from 2007 to 2010)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AARE</td>
<td>117.09</td>
<td>118.85</td>
<td>107.10</td>
<td>76.25</td>
<td>68.63</td>
</tr>
<tr>
<td>RMSE</td>
<td>85.96</td>
<td>67.98</td>
<td>68.93</td>
<td>67.92</td>
<td>48.46</td>
</tr>
<tr>
<td>PRC</td>
<td>0.7269</td>
<td>0.7540</td>
<td>0.7518</td>
<td>0.7936</td>
<td>0.8944</td>
</tr>
</tbody>
</table>

the forecasting results of five different models for 48 testing samples, we can see that the forecasting results of M–RBF–NN model are the best in all models, and the M–RMF–NN can better capture the mapping relation than the other four model.

Figure 4–8 show the forecasting results of five different models for 48 testing samples. We can see that the forecasting results of M–RBF–NN model are best out of all five models. From the graphs and table, we can generally see that the forecasting results are very promising in the rainfall forecasting under the research where either the measurement of fitting performance is goodness or where the forecasting performance is effectiveness. It can also be seen that there was consistency in the results obtained between the training and testing of these M–RBF–NN model.

In comparison of model AR(5) with model MLP–NN, both of which used the same input data, the model MLP–NN yielded better results than model AR(5) for both the training and testing samples. The results show the rainfall system is a complex nonlinear system and the traditional statistical model is very difficult to use for accuracy prediction.

For model MLP–NN and S–RBF–NN, the results of these two models are closer in testing samples. As shown in Table 2, the results show the RMSE of the MLP–NN model is 67.98 and the RMSE of the S–RBF–NN model is 68.93 about the rainfall forecasting. Similarly, PRC of the MLP–NN model is 0.7540 and the PRC of the S–RBF–NN model is 0.7518. Figure 5 and 6 reveal that both models MLP–NN and S–RBF–NN provided underestimated rainfall.
forecasts, showing better performance than Model AR(5). Models MLP–NN and S–RBF–NN are based on neural network theory, but those algorithms are different. Those results indicate that model neural network is capable of modelling without prescribing hydrological processes, catching the complex nonlinear relation of input and output, and solving without the use of differential equations.

As shown in Table 2 for model SR, remarkable performance indicates that model SR is capable of generalizing better results from the same set of input variables than model AR(5), MLP–NN and S–RBF–NN. The results of the modular model can significantly improve the prediction accuracy. Model M–RBF–NN, which involved the same input data of rainfall at the Liuzhou, produced the highest performance. For example, the AARE of the M–RBF–NN is 68.63, the RMSE of the M–RBF–NN model is 48.46, and the PRC of the M–RBF–NN model is 0.94. The values of AARE and RMSE are the minimum and the values of PRC are the maximum in all models. The results indicate that the deviations between original values and forecast are very small, and the modular model is capable of capturing the average change tendency of the daily rainfall data.

From the experiments presented in this study we can draw that the M-RBF-NN model is superior to other models in fitting and testing cases in terms of the different measurement, as can be seen in Table 2. There are three main reasons for this phenomenon. Firstly, the rainfall
system contain complex nonlinear pattern. SSA can extract complex trends and find structure in rainfall time series. Using a different the kernel function form of RBF can establish the effective nonlinear mapping for rainfall forecasting. Secondly, the output of different models has the high correlative relationship, the high noise, nonlinearity and complex factors. If PLS technology doesn’t reduce the dimension of the data and extract the main features, the results of the model will be unstable. At last, LS–SVR is used to combine the selected individual forecasting results into a nonlinear ensemble model, which keeps the flexibility of the nonlinear model. Therefore the proposed nonlinear modular ensemble model can be used as a feasible approach to rainfall forecasting.

4. Conclusion

Accurate rainfall forecasting is crucial for a frequent unanticipated flash flood region to avoid losing lives and economic loses. In this study, modular Radial Basis Function Neural Network model was employed to forecast monthly rainfall for Liuzhou, Guangxi. In terms of different forecasting models, empirical results show that the developed modular model performs the best in prediction monthly rainfall on the basis of different criteria. Our experimental results demonstrated the successful application of our proposed new model, M–RBF–NN, for the complex forecasting problem. It demonstrated that it increased the rainfall forecasting accuracy more than any other model employed in this study in terms of the same measurements. Therefore, the M–RBF–NN ensemble forecasting model can be used as an alternative tool for monthly rainfall forecasting to obtain greater forecasting accuracy and improve the prediction quality further in view of empirical results, and can provide more useful information, and avoid invalid information for the future forecasting.

Acknowledgment

The authors would like to express their sincere thanks to the editor and anonymous reviewers comments and suggestions for the improvement of this paper. This work was supported by Program for Excellent Talents in Guangxi Higher Education Institutions, by Natural Science Foundation of Guangxi under Grant No. 2011GXNSFE018006 and by the Natural Science Foundation of China under Grant No.11161029.

References


