Analysis of rainfall depth based on climatology conditions using artificial neural networks

To cite this article: Very Dermawan and Yuzy Alfahnie 2020 IOP Conf. Ser.: Earth Environ. Sci. 437 012020

View the article online for updates and enhancements.
Analysis of rainfall depth based on climatology conditions using artificial neural networks

Very Dermawan¹, Yuzy Alfahnie²

¹ Water Resources Engineering Department, Faculty of Engineering, Universitas Brawijaya Malang, 65145, East Java - Indonesia.
² Water Resources Engineering Department, Faculty of Engineering, Universitas Brawijaya Malang, 65145, East Java - Indonesia.
E-mail: peryderma@ub.ac.id

Abstract. The quantity of rain that falls on the earth cannot be known with certainty. Floods and droughts due to a small quantity of rainfall are frequent events in some areas of Indonesia. The depth of rainfall at a certain time can be anticipated with accurate information. Along with rapid advances in technology, the forecasting of patterns of rainfall can be performed by artificial intelligence models, using historical data for the climatological parameters. The aim of this study is to predict rainfall depth based on climatology data. There are three categories of data that were obtained using NeuroSolutions for Excel: monthly, daily, and hourly data. The input data are temperature, pressure, duration of sunshine, and humidity. The output data is rainfall depth. Based on the results of running calculations on monthly, daily, and hourly data, it was indicated that monthly, daily, and hourly data have relative errors of 11.49%, 8.49%, and 19.32% respectively.

Keywords: Rainfall depth, Climatology, Artificial Neural Network

1. Introduction
Forecasting of weather is one of the most imperative and demanding operational responsibilities carried out by meteorological services [1, 2]. It is complicated due to uncertainty. Rainfall is a natural climatic phenomenon for which its prediction is challenging and demanding [3, 4]. Rainfall is one of the most eminent and complex atmospheric phenomena. It is complex as it involves interaction of several atmospheric processes and is vital for survival and sustenance of the earth [5]. Precipitation prediction is very important for countries whose economies depend mainly on the agriculture sector.

Rainfall plays a very important role in agriculture in that it is very essential for irrigation. Rainfall forecasting is also very useful for water management, flood forecasting, hydroelectric power generation, drinking water supply, reservoir operation, and urban drainage planning [6, 7, 8, 9]. Many factors of climatology data will influence rainfall prediction, such as temperature, humidity, wind speed, pressure, dew point, duration of sunshine, and so on.
Artificial Neural Network (ANN) models have been widely used in various models of hydrological processes with high accuracy. The main aim of the present work is to construct an ANN model of rainfall forecasting patterns using historical data regarding climatological parameters. In this study, ANN models have been identified for prediction of rainfall depth.

Indonesia, which is a developing country in Southeast Asia and where a majority live through agriculture activities, has problems in providing these hydrological data. The availability of rain gauges, river water level recorders, and weather and climate parameter measurement devices is not even and well distributed in each catchment area. Even in some watersheds, there are no rain gauges or water level recorder instruments. In some watersheds that have measuring gauges for rainfall and river water levels, sometimes there are problems in maintenance, and thus much of the data is not recorded due to gauge damage. In watersheds in Indonesia, hydrological data, especially rainfall data, are commonly measured by a manual daily rainfall gauge; there are not many automatic rainfall recorders that are set to measure hourly data. More often, daily data is also not recorded continuously because the measuring instrument is broken.

This study is conducted to give solutions for this hydrological data problem, especially in prediction of missing rainfall data at certain times. In this study, rain forecasting analysis was carried out using climatological (weather) data. This is because in Indonesia, the recording of climatological (weather) data is better and more accurate than the recording of rainfall data in certain regions. This study is carried out with the assumption that it will be useful for forecasting the missing data, especially for daily data. Once known with sufficient accuracy, the method offered will be used in analyses to estimate the amount of monthly, daily, and hourly rainfall by analysing weather data sets.

2. Material and Methods
The data set used in this study was generated through continuous monitoring of the climatologic data of the City of Malang, East Java, Indonesia, such as rainfall depth, temperature, pressure, humidity, and duration of sunshine. This research was conducted on climate monitoring stations located in the City of Malang, as the Karangploso and University of Brawijaya rain gauge stations. These were the utilized data of climatologic parameters:
(a) Rainfall data, consisting of monthly rainfall data over 25 years (1987-2011), daily rainfall data over 6 years (2008-2013), and hourly rainfall data over 5 years (2009-2013), and
(b) Weather data (temperature, pressure, duration of sunshine, and humidity), consisting of monthly data over 25 years (1987-2011), daily data over 6 years (2008-2013), and hourly data over 5 years (2009-2013). Analysis of rainfall depth prediction was carried out by an artificial neural network model using NeuroSolutions for Excel. NeuroSolutions for Excel is a Microsoft Excel add-in that simplifies and enhances the process of getting data into and out of a NeuroSolutions neural network. This tool benefits both novice and advanced neural network developers by offering ease of use, yet extremely powerful features. The foremost feature of this product is that all tasks can be performed directly from Excel.

NeuroSolutions uses a single layer perceptron, which unfolds a dynamic net at each time step into an equivalent feed-forward net. The “Analyze Data” module provides the user with useful information about the data. The operations available in this module can be used during the pre-processing stage of neural network design or to analyse the network output. The “Tag Data” module provides a simple graphical method for tagging portions of data as Training Input, Training Desired, Cross Validation Input, Cross Validation Desired, Testing Input, Testing Desired, and Production Input. The “Train Network” module gives the user the ability to train a network once, multiple times with different random initial conditions, or multiple times while varying one or more network parameters.
In this study, weather data such as temperature, pressure, duration of sunshine, and humidity were utilized for input parameters, and rainfall depth was utilized for output desired with an epoch of 1000. For running the model, the data composition is listed below:


b. Daily Data: 3 years (2008-2010) as training, 1.5 years (January 1, 2011 to June 30, 2012) as cross-validation, 3 years (2007-2009) as testing, 6 months (July 1, 2013 to December 30, 2013) as production, and a period in 2012 (March 29 to April 26) as data production.

c. Hourly Data: 2 periods from 2009 (April 2 - April 17) to 2010 (March 29 - April 13) as training, 12 data sets in 2011 (March 14 - March 29) as cross-validation, 12 data sets in 2011 (March 30 - April 14) as testing, 6 months in 2012 (July 1 - December 30) as production, 1 month in 2012 (March 29 – April 26) as data production.

3. Result and Discussion

In this study, the ANN method was applied to predict rainfall depth for the City of Malang. The prediction stage of the ANN method in this study was made of three dataset configurations composed of monthly data, daily data, and hourly data. The differences between the actual rainfall data and output from the ANN method are shown by the average relative error.
Figure 3. Relationship of MSE and Epoch for daily data

Figure 4. Relationship of MSE and Epoch for hourly data

Figure 5. Desired output and actual network output from monthly data sets (average relative error = 11.49%)
Figure 6. Comparison of ANN output and actual output from monthly data sets (average relative error = 11.49%)

Figure 7. Desired output and actual network output from daily data sets (average relative error = 8.49%)

Figure 8. Comparison of rainfall depth from ANN output and actual depth from daily data sets (average relative error = 8.49%)
Figure 9. Desired output and actual network output from hourly data sets (average relative error=19.32%)

Figure 10. Comparison of rainfall depth from ANN output and actual output from hourly data sets (average relative error=19.32%)

The general equation of rainfall depth modelling with ANN consists of B as bias and X as the description of climatological elements, where $X_1$ is temperature, $X_2$ is humidity, $X_3$ is duration of sunshine, and $X_4$ is air pressure. Z is a hidden network, and W is the weight of the network. The ANN output of monthly rainfall modelling in 1987-2011 based on monthly climatologic data at Karangploso climatology station had an average relative error of 11.49%. Below is the general equation of ANN modelling based on monthly data sets:

$$y = W_x + \sum_{k=1}^{n} \left( b_{ij} + \left( \sum_{j=1}^{m} W_{xj} \right) \right) + \left( b_{kl} + \left( \sum_{m=1}^{n} W_{zm} \right) \right) + \left( b_{lj} + \left( \sum_{n=1}^{m} W_{ln} \right) \right)$$

The general equation of ANN modelling based on daily data sets consists of B as bias and X as the description of climatological elements, where $X_1$ is temperature, $X_2$ is humidity, and $X_3$ is duration of sunshine. The element $X_4$ as pressure was eliminated. Modelling of daily rainfall depth in 2008-2013 based on daily climatologic data at Karangploso climatology station had a relative error of 8.49%. Based on daily rainfall data sets, below is the general equation of ANN modelling:
The general equation of rainfall depth modelling with ANN based on hourly data sets consists of $B$ as bias and $X$ as the description of climatological elements, where $X_1$ is temperature, $X_2$ is humidity, $X_3$ is duration of sunshine, and $X_4$ is air pressure. Modelling of hourly rainfall in 2012 based on hourly data at the Hydrology Laboratory station of the University of Brawijaya had a relative error value of 19.32%. Below is the general equation of ANN modelling based on hourly data sets:

$$ Y_k = W_0 + \sum_{j=1}^{n} \left( B_{1j} + \left( \sum_{i=1}^{3} X_i Z_{ij} + X_4 Z_{ij} \right) W_{1j} \right) + \left( B_{2j} + \left( \sum_{i=1}^{3} X_i Z_{ij} + X_4 Z_{ij} \right) W_{2j} \right) $$

4. Conclusion
This study was able to give solutions for prediction of missing rainfall data at certain times using climatological or weather data. Due to the accuracy of climatological or weather data in certain regions, this study will be useful for forecasting the missing data.

Based on the results of running the ANN model, it can be seen that the best prediction of the rainfall depth is based on daily climatologic data sets. This means that the best prediction of rainfall depth for the City of Malang is based on daily weather data.

References