Chapter 11
Reconstruction of Missing Hourly Precipitation Data to Increase Training Data Set for ANN

Hema Nagaraja
Department of Computer Science, Jaypee Institute of Information Technology, Noida, India

Krishna Kant
Manav Rachna International University, Faridabad, India

K. Rajalakshmi
Jaypee Institute of Information Technology, Noida, India

ABSTRACT

This paper investigates the hourly precipitation estimation capacities of ANN using raw data and reconstructed data using proposed Precipitation Sliding Window Period (PSWP) method. The precipitation data from 11 Automatic Weather Station (AWS) of Delhi has been obtained from Jan 2015 to Feb 2016. The proposed PSWP method uses both time and space dimension to fill the missing precipitation values. Hourly precipitation follows patterns in particular period along with its neighbor stations. Based on these patterns of precipitation, Local Cluster Sliding Window Period (LCSWP) and Global Cluster Sliding Window Period (GCSWP) are defined for single AWS and all AWSs respectively. Further, GCSWP period is classified into four different categories to fill the missing precipitation data based on patterns followed in it. The experimental results indicate that ANN trained with reconstructed data has better estimation results than the ANN trained with raw data. The average RMSE for ANN trained with raw data is 0.44 and while that for neural network trained with reconstructed data is 0.34.

DOI: 10.4018/978-1-6684-2408-7.ch011

Copyright © 2022, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

1. INTRODUCTION

Meteorological data are a set of climatic information, which describes the characteristics of the atmosphere. Meteorological station records climatic parameters like air temperature, dew point temperature, atmospheric pressure, rainfall, wind speed, wind direction, maximum temperature, minimum temperature and sunshine hours in minutes. These data are available on the hourly or daily basis. Meteorological data are very important in hydrological analysis and for agricultural purpose. Measuring instrument or infrastructural failures creates missing data in the measurements. These missing data may be significant portions and can be random in space and time. This paper proposes to reconstruct the missing hourly rainfall and how this increased data set will help in improving estimation of precipitation.

Water reservoir system, flood warning systems, domestic usage planning, irrigation calculation, industry usage planning and planning of hydropower generation (Mays, 2010) required accurate hourly rainfall data. Real-time smart irrigation system requires accurate and complete hourly rainfall data from past and present hour to schedule the timely and right amount of irrigation as described in "Specification for Weather-Based Irrigation Controllers" by WaterSense (July, 2012). Incomplete rainfall data provided to real-time irrigation system will lead to over or under irrigation for that particular irrigation cycle.

Filling of missing values can be based on the same site or use nearby sites to estimate values. Estimation is based on time dimension, space dimension or a combination of both depending on the pattern of missing data and its correlation with each other (Gao et al., 2015). Higher the correlation better estimation of missing values (Graham, 2012). The selection of interpolation method mainly depends on the type of data that are used to fill the gaps. Rainfall data are spatiotemporal variable and uses deterministic and geostatistical methods for estimation of missing rainfall data (Piazza et al., 2015). Understanding of precipitation patterns is important for appropriate selection of interpolation method.

A study by Deshpande et al. (2012) provides the characteristics of the hourly rainfall over India based on a large network of Self Recording Rain Gauge Stations with 25 years of data. Study shows that area of our interest Delhi, India has rainfall of minimum 200 hr/yr and a maximum of 600 hr/yr. During non-monsoon season that is for January to May and October to December the average rainfall is less than 10 cm and during monsoon season average rainfall is between 50 cm to 100 cm. Characteristics of acquired data from Delhi validates the characteristics defined by Deshpande et al. (2012).

For experimentation purpose, real data has acquired from 14th January 2015 to 28th February 2016 from eleven Automatic Weather Stations (AWS) located in Delhi. Approximate annual rainfall for all eleven AWS has summarized in Table 1. Column 1 of Table 1 is a total number of rainfall hours for each of eleven AWS's for the year 2015 to 2016. The average rainfall is 577 hr, and is falling within the range defined by Deshpande et al. (2012). Column 2 of Table 1 has a total number of hours of rainfall occurred in a non-monsoon season of the year 2015-16. On average 172 hours of rainfall has occurred in a non-monsoon season, with average rainfall range of 1-9 mm/hr. Monsoon has frequent rainfall with an average of 406 hr with a range of 1-55 mm/hr.

Rain gauges used in AWS is of tipping bucket type. The range of the sensor is the 0 to the 1023mm/hr with an accuracy of ±5%. Rain gauge's recorded data is hourly cumulative and refreshed at 3:00 UTC. Indian Meteorological Department (IMD) Delhi (semi-arid region) data from eleven AWS's has following hourly rainfall characteristics:

- Most of the time there is no precipitation and data read from the meteorological station are 0 mm.
- Single AWS showing rainfall has a low intensity of precipitation in the range of 1-6 mm/hr.

22 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/reconstruction-of-missing-hourly-precipitation-data-to-increase-training-data-set-for-ann/288959

Related Content

Mapping Ground Penetrating Radar Amplitudes Using Artificial Neural Network and Multiple Regression Analysis Methods
Eslam Mohammed Abdelkader, Mohamed Marzoukand Tarek Zayed (2022). *Research Anthology on Artificial Neural Network Applications (pp. 1052-1076).*
www.irma-international.org/chapter/mapping-ground-penetrating-radar-amplitudes-using-artificial-neural-network-and-multiple-regression-analysis-methods/288999

Cooperative Control of Unknown Networked Lagrange Systems using Higher Order Neural Networks
Gang Chenand Frank L. Lewis (2013). *Artificial Higher Order Neural Networks for Modeling and Simulation (pp. 214-236).*
www.irma-international.org/chapter/cooperative-control-unknown-networked-lagrange/71801

A Method of Estimation for Magnetic Resonance Spectroscopy Using Complex-Valued Neural Networks
Naoyuki Morita (2009). *Complex-Valued Neural Networks: Utilizing High-Dimensional Parameters (pp. 256-283).*
www.irma-international.org/chapter/method-estimation-magnetic-resonance-spectroscopy/6772

Study and Analysis of Visual Saliency Applications Using Graph Neural Networks
Gayathri Dharaand Ravi Kant Kumar (2023). *Concepts and Techniques of Graph Neural Networks (pp. 108-131).*
www.irma-international.org/chapter/study-and-analysis-of-visual-saliency-applications-using-graph-neural-networks/323825

A Detailed Study on Classification Algorithms in Big Data
Saranya N.and Saravana Selvam (2020). *Big Data Analytics for Sustainable Computing (pp. 30-46).*
www.irma-international.org/chapter/a-detailed-study-on-classification-algorithms-in-big-data/238603