Desert Ecosystem Engineering Journal

Journal homepage: http://deej.kashanu.ac.ir

University of Kashan
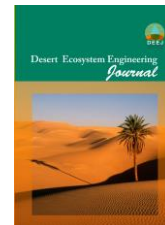
Use of global climate model outputs to fill missing temperature and precipitation missing data (Case study: Khanmizza Basin Watersheds)

Sara, Mardanian[1], Abdollahi Khodayar [2], Rafat Zare Bidaki[3], Abozar Salehi[4]

Extended abstract

Introduction: Due to importance of data quality, issues relating to filling the missing data has found a great deal of interest. Regeneration methods for missing data can be classified into two kinds of classical and modern categories. Application of statistical methods such as relationship with nearby stations and approaches on the base of hydrological, climatological or physiographical similarity are among the common techniques for gap filling. For such cases, correlation with elevation is one of the common techniques. The modern methods make use of approaches such as data mining, ANN, fuzzy arithmetic and so on to improve the completeness of data. Typically, the output of climate models is used for future predictions/assessment. After development of climate model like general circulation model (GCM) several researchers have reported a good correlation between global atmosphere or ocean and local station data. This correlation has a potential usage for estimation of missing data at local scale.

Materials and methods: Two types of data, including meteorological data and climatic data outputs, were used in this study. Reconstruction of two key variables at monthly scale, temperature and precipitation, were investigated over a period of 20 years (1995-2016). Climate data for these two climatic parameters downloaded from Canadian website site (GCM / RCM), were used as an auxiliary variable for data reconstruction. The least squares error method was applied to determine the scenario with the most realistic site-based conditions. The coefficient of determination was used as a measure to assess how well scenarios was able predicts observed data. For this purpose, four scenarios that generally showed either a high linear or non-linear correlation were selected. Then, by means of cross validation and the evaluating relationship between these two series of data and observation were tested. The outcome at this stage was 4 out of climate scenarios which show a high correlation with local station. Cross validation was employed the final scenario. Here the correlation between climate data and station data was utilized as an auxiliary variable for reconstruction of missing values. The outputs of new method were compared with elevation-based method.

1. Graduate Student Department of Rangeland and Watershed Engineering, Shahrekord University, Kabdolla2010@gmail.com
2. Assistant Professor Shahrekord University, Kabdolla2010@gmail.com
3. Assistant Professor Shahrekord University
4. Graduated Master of Water Resources Engineering

Results: The established relationship between observations and climatic data (climate model output) by the least squares error method showed a high correlation for the selected models, thus ECHO-G and NCARCCSM3 models were selected for rainfall and temperature respectively. The performance of these predictive models was tested by evaluating them after assuming each row data as missing data. The observational data was re-organized into monthly, seasonal and annual time interval. The obtained coefficients were successfully verified against observed data by eliminating the data for each month and establishing a new relationship. Four scenarios including ECHO-G (SR-A1B), ECHO-G (SR-A2), NCARCCSM3 (SR-A2) and NCARCCSM3 (SR-B1), showed a high correlation between climatic and calculated data. Although the correlation coefficient between rainfall and the climatic model output was low but a good correlation between the combination factor of rainfall and temperature with observed data was found in the study area (r=0.98). The established relationships for local station even can be used for future (up to 2050) predictions. A comparison of the developed method with the elevation-based method showed a better performance at monthly scale.

Discussion and conclusion: Awareness of climatic conditions is a helpful strategy to reduce the damage caused by natural disasters. This most often requires an extensive and continuous data sets. Traditionally, correlation with elevation was used as an auxiliary variable to fill the gaps in rainfall and temperature data. Previous studies have reported a strong correlation between the local data of stations and climatic data. This study was conducted to investigate the possibility of using downloaded data from Canadian climate site (GCM / RCM for filling temperature and precipitation meteorological data. The results of the new method were more consistent with observed data. Cross validation indicates the applicability of using climate model as an auxiliary variable for regenerating missing rainfall and temperature data or predicting these variables in the area of interest. In addition to fill the missing data, the common advantage the use of climate model output is that, it could be used as future predictor. On the other hand, the comparison of the above mentioned method with the elevation-based method with observational data indicated that the new approach provide a solution especially for the case that there exist no or limited number of neighboring stations.

Keywords: Global climate model, Climate change, Khanmyrza Basin, Least square error and Canadian climate