PRECIPITATION DATA MERGING USING GENERAL LINEAR REGRESSION

Anish C. Turlapaty[1, 2] , Nicolas H. Younan[1,2] , Valentine Anantharaj[2]
[1]Department of Electrical and Computer Engineering, Mississippi State University, Mississippi State, MS 39762
[2]Geosystems Research Institute, Mississippi State University, Mississippi State, MS 39762, USA

Introduction: Precipitation is a key process and has a significant role in global energy cycle. It is the process of great interest in weather forecast models and climate studies. Precipitation is the major force behind the natural disasters like floods and storms. Currently there are many satellite based rainfall products like PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks) provide a spatial resolution of 0.25 by 0.25 deg. However, there are no satellite measurements providing global measurements with high spatial and temporal resolutions. National Aeronautics and Space Agency (NASA) researchers proposed a constellation of satellites known as global precipitation measurement mission, due for launch in a few years. The objective of this mission is to obtain global precipitation data at very high temporal and spatial resolution. These global measurements will benefit the researchers studying both short term and long term meteorological phenomena, especially in understand the spatio temporal signals in the precipitation data on a global scale and thus better relate the rainfall with other climate processes. Moreover, these global datasets are useful for those countries with very little ground based measurement equipment. For instance, the GPM measurements are useful for nations like Bangladesh where floods and storms are major issues as they do not have enough rain gage networks to study the floods.

Methodology: The Geosystems Research Institute (GRI) at Mississippi State University received funding from NASA to perform a GPM rapid prototyping capability experiment. The objectives are to develop an intelligent merging methodology of different observation sets, develop a spatially downscaled product, evaluate the downscaled product, and to estimate the uncertainties of the merged products. The GRI researchers used the precipitation observations available from several satellites and ground based sensors to develop downscaled spatio-temporal data on a uniform grid with a spatial resolution of 0.1º by 0.1º and a time resolution of 1 hour or less. Toward partial fulfillment of the project objectives, a data fusion methodology is proposed to merge the downscaled precipitation datasets to develop a merged product which is statistically superior to any individual data set or their average. The datasets used in this work are 1) precipitation data from Climate Prediction Center Morphing method (CMORPH) [1]; 2) data from Auto Estimator algorithm for Geostationary Operational Environmental Satellite data (GOES AE) [2]; 3) data from Hydro estimator algorithm for Geostationary Operational Environmental Satellite data [3]; 4) data from A Self Calibrating Real-Time GOES Rainfall Algorithm (SCAMPR) [4]. The test region is a rectangular grid of 80 by 200 cells surrounding the region of Arkansas Red Basin River Forecast Center. A General Linear Regression model is applied to Spatio-temporal precipitation data from different products and a reference data available from the Arkansas basin (ABRFC) data. The regression is performed between the time-series of reference data on a 3 x 3 spatial grid and four sets of time-series of over the same spatial grid. Thus the reference time-series is the regressand and the individual precipitation data are regressors. The general regression results in a regression coefficient vector which relates the individual observations to the reference data. The assumption in this method is that a similar regression relation exists among the individual datasets and the actual rainfall and a spatial correlation exists between rainfall time-series at a grid cell and its neighboring cells. In order to improve the accuracy of the merged product and using the spatial correlation between neighbors, a linear propagating error model is developed using simple linear regression between the error vectors of neighboring grid cells [5]. The gradient obtained from this model is used to develop a correction vector at the grid cell of interest. By recursive application of this method along the spatial extent of the datasets, a complete merged product is developed.

Results and Conclusion: This resulting dataset is evaluated against the individual datasets and the average dataset using simple statistical measures like bias, correlation, efficiency, and mean square error to mention a few. The merged product is also evaluated in terms of metrics like false detection rate and successful detection rate. The merged dataset is statistically better than other sets. For instance, in general, the false alarm rate of the average data is 80 to 90% whereas for the merged data it is 50 to 40% as shown in figure 1. Thus, from the initial results, the general regression based fusion methodology promises to be a good choice for merging precipitation measurements if reference data is available at a few spatial locations.

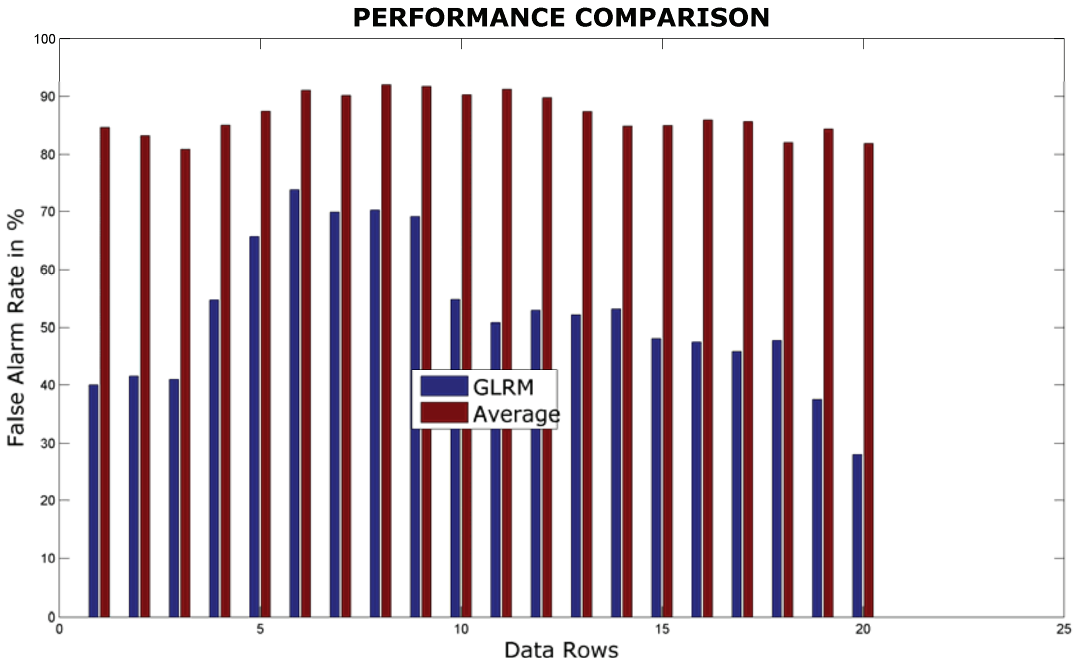

Figure 1: Comparison of false alarm rate in % of the merged data with average precipitation data at 20 sample rows. X-axis is data rows, Y-axis is percentage false alarm rate, Blue bars: GLRM, Red: Average

References

[1] R. J. Joyce, J. E. Janowiak, P. A. Arkin, and P. Xie, "CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution", Journal of Hydrometeorology. Vol 5, pp 487 – 503. 2004

[2] G. A. Vincente, R. A. Scofield, W. P. Menzel, "The Operational GOES Infrared Rainfall Estimation Technique", Bulletin of the American Meteorological Society. vol:79, no.9,pp1883–1898. 1998

[3] R.A. Scofield, R.J.Kuligowski, "Status and Outlook of Operational Satellite Precipitation Algorithms for Extreme-Precipitation Events",Weather & Forecasting; vol:18, no.6, pp1037-1051, 2003.

[4] R. J. Kuligowski, "A Self Calibrating Real-Time GOES Rainfall Algorithm for Short-Term Rainfall Estimates", Journal of Hydrometeorology, vol:3, no.2. pp112-130, 2002.

[5] D. Pokrajac and Z. Obradovic, "Combining Regressive and Auto-Regressive Models for Spatio-Temporal Prediction", Proc. 17th International Machine learning workshop on spatial knowledge, june 2000.