

EGU21-14650

<https://doi.org/10.5194/egusphere-egu21-14650>

EGU General Assembly 2021

© Author(s) 2021. This work is distributed under the Creative Commons Attribution 4.0 License.



## K nearest neighbour in merging satellite rainfall estimates from diverse sources in sparsely gauged basins

**Biswa Bhattacharya**<sup>1</sup> and Junaid Ahmad<sup>2</sup>

<sup>1</sup>IHE Delft Institute for Water Education, Hydroinformatics and Socio-Technical Innovation, Delft, Netherlands

([b.bhattacharya@un-ihe.org](mailto:b.bhattacharya@un-ihe.org))

<sup>2</sup>Punjab Irrigation Department, Lahore, Pakistan

Satellite based rainfall estimates (SBRE) are used as an alternative to gauge rainfall in hydrological studies particularly for basins with data issues. However, these data products exhibit errors which cannot always be corrected by bias correction methods such as Ratio Bias Correction (RBC). Data fusion or data merging can be a potentially good approach in merging various satellite rainfall products to obtain a fused dataset, which can benefit from all the data sources and may minimise the error in rainfall estimates. Data merging methods which are commonly applied in meteorology and hydrology are: Arithmetic merging method (AMM), Inverse error squared weighting (IESW) and Error variance (EV). Among these methods EV is popular, which merges can be used to merge bias corrected SBREs using the minimisation of variance principle.

In this research we investigated the possibility of using K nearest neighbour as a data merging method. Four satellite rainfall products were used in this study namely CMORPH, PERSIANN CDR, TRMM 3B42 and MSWEP. MSWEP was used as a reference dataset for comparing the merged rainfall dataset since it is also a merged product. All these products were downloaded at 0.25° x 0.25° spatial scale and daily temporal scale. Satellite products are known to behave differently at different temporal and spatial scales. Based on the climatic and physiographic features the Indus basin was divided into four zones. Rainfall products were compared at daily, weekly, fortnightly, monthly and seasonal scales whereas spatial scales were gauge location, zonal scales and basin scale. The RBC method was used to correct the biasness of satellite products by correcting the products at monthly and seasonal scale. With bias correction the daily normalised root mean square error (NRMSE) was reduced up to 20% for CMORPH, 22% for PERSIANN CDR and 14% for TRMM at the Indus basin scale for monthly scale which is why the monthly bias corrected data was used for merging. Merging of satellite products can be fruitful to benefit from the strength of each product and minimize the weakness of products. Four different merging methods i.e. Arithmetic merging method (AMM), Inverse error squared weighting (IESW), Error variance (EV) and K Nearest Neighbour method (KNN) were used and performance was checked in two seasons i.e. non-wet and wet season. AMM and EV methods performed similarly whereas IESW performed poorly at zonal scales. KNN merging method outperformed all other merging methods and gave lowest error across the basin. Daily NRMSE was reduced to 0.3 at Indus basin scale with KNN method, AMM and EV reduced the error to 0.45 in comparison to error produced by CMORPH, PERSIANN

CDR and TRMM of 0.8, 0.65 and 0.53 respectively in the wet season. KNN merged product gave lowest error at daily scale in calibration and validation period which justifies that merging improves rainfall estimates in sparsely gauged basin.

**Key words:** Merging, data fusion, K nearest neighbour, KNN, error variance, Indus.