Streamflow forecasting using an integrated methodology based on rainfall-runoff modeling, precipitation ensemble predictions and machine learning techniques

Luz Adriana Gómez[1], Carlos D. Hoyos[1], Nactaly Posada[1], Diana Carolina Cruz[1], Juan José Toro[1], Juan Esteban Taborda[1] and David Muñoz[1]

(1) Department of Geosciences and Environment, Universidad Nacional de Colombia, Sede Medellín, Colombia

Streamflow forecasting is a vital tool for water resources management and represents a fundamental task for successful hydropower operation. In that sense, it is important to advance in the optimization and accuracy of operational forecasting systems, especially for high variable daily streamflow in tropical-mountainous basins. In this study, a three-stage forecasting system is proposed. In the first stage, we deal with one of the important issues in hydrological modeling, the determination of optimal initial conditions. For that, we designed a continuous scheme based on the water storage obtained from hydrological simulations of previous days and having bias-corrected satellite information from IMERG as input data. In the second stage, those initial conditions along with bias-corrected ensemble forecasts of the European Centre for Medium-Range Weather Forecasts (ECMWF) for a 15-day lead-time are used to drive a calibrated distributed rainfall-runoff model. Finally, considering past errors during the forecasting process, the third stage of the methodology comprises post-processing using a quantile mapping bias correction method and an output updating procedure based on machine learning. For the latter, we compare the model performance improvement using two statistical techniques: multiple linear regression and K nearest neighbors (KNN) algorithm. The methodology was evaluated through application to the Sogamoso river basin located in a complex terrain over the northern part of the Andes cordillera. Results indicate that the proposed forecasting system allowed a reliable streamflow forecast in terms of variability and magnitude, especially for 1-5 days ahead. Besides, the post-processing manipulations allowed to reduce forecast bias problems during different states of streamflow seasonality. Temporal changes in the forecast performance were assessed over 30-day rolling windows using correlation, bias, root mean square error, and continuous ranked probability skill score (CRPSS).