



Increasing complexity in model structure and likelihood function helps to identify dominant streamflow mechanisms: A case study of two small forest basins in Brazil

Paula Cunha David, Debora Yumi de Oliveira, and Pedro Luiz Borges Chaffe

Federal University of Santa Catarina, Department of Sanitary and Environmental Engineering, Brazil
(paulacunhadavid@gmail.com)

Despite the fact that even similar basins might hide different dominant runoff mechanisms, many conceptual hydrological models usually have a fixed structure. In addition, model calibration using likelihood functions whose assumptions are not satisfied or checked leads to unreliable parameter and uncertainty estimates. The objectives of this work were twofold: (i) to evaluate the impact of the use of different likelihood functions in model calibration; and (ii) to identify the model structures that better represent the rainfall-runoff process in two forest basins located in the southern region of Brazil. The models were calibrated with the automatic calibration algorithm DREAM(ZS), using three likelihood functions with increasing complexity: the first one considers that the errors are Gaussian and independent; the second one considers the heteroscedasticity of the residuals; and the third one considers a non-normal distribution for residuals, in addition to heteroscedasticity. Ten different structures from the SUPERFLEX framework were tested, varying the number of reservoirs and parameters, inclusion of lag functions and non-linearity of the reservoirs. The results were evaluated using three metrics: reliability, precision and volumetric bias. Two information criteria were used to rank the models: the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). For the Saci River Basin (0,1 km²), the more complex likelihood function best met the assumptions of the residual model and presented a better predictive distribution. In the Bugres River Basin (6,94 km²), the third likelihood function presented worst results, due to the correlation of the parameters of the residual model and the hydrological models. When evaluated with the three likelihood functions, the information criteria resulted in different model ranking for each basin. Considering both information criteria and quality of the predictive distribution, the use of an unsaturated reservoir followed by two independent reservoirs represented the basin in a better way for the Saci River Basin, while for the Bugres River Basin the models that presented the best result are nonlinear, which may indicate that the basin has a significant nonlinear behavior. Adding complexity in a systematic way helped to identify the most important components of the model structures. The use of information criteria allowed to choose the model that better represents each basin, considering both parsimony and goodness of fit. The assumptions about model residuals were not totally satisfied with any of the likelihood functions. Future studies are needed to identify a likelihood function whose assumptions are completely satisfied, which is important to correctly choose the best model.