

Comparison of natural streamflows generated from a parametric and nonparametric stochastic model

James Prairie(1,2), Balaji Rajagopalan(2) and Terry Fulp(1)

1. CADSWES, University of Colorado at Boulder

2. Dept. of Civil, Env. and Arch. Engg., University of Colorado at Boulder

Abstract. Stochastic techniques for simulating synthetic streamflow sequences that are "statistically faithful" to the historical data, are important for various water and environmental decision making situations. Parametric models such as Autoregressive (AR) and Autoregressive Moving Average (ARMA) models are widely used in practice. Parametric models assume Gaussian distribution of the data - if not appropriate, transformation of the data has to be done prior to model fitting. Consequently, they are limited in their ability to reproduce non-Gaussian (e.g. skewness, bimodality) and nonlinear (e.g. nonlinearity in the month to month flows) structures that the data might exhibit. Furthermore, the parametric models require estimation of several parameters, which increases with the order of the model - often from limited data. Parameter estimation issues and their variability due to small sample size, further adds to the limitations. Nonparametric techniques alleviate some of the drawbacks of parametric models and also provide an assumption free framework. They have the ability to reproduce any arbitrary structure present in the data and are easily portable across sites. Several nonparametric techniques have been developed in recent years. In this study we use the K-NN (K-nearest neighbor) technique. The technique involves generating flows for subsequent time steps from a set of K "similar patterns" (or nearest neighbors) identified from the historical data, to the pattern at the current time step. We offer comparisons of K-NN technique and parametric model as developed in the package Stochastic Analysis, Modeling, and Simulation (SAMS) on synthetic data and also on the streamflow data from USGS stream gauge 09180500: Colorado River near Cisco, Utah. The simulations from these two methods will be compared on the reproduction of various distributional attributes of monthly and annual flows (e.g. mean, variance, co-efficient of skewness, probability density function). Other attributes such as bivariate and conditional pdfs of consecutive monthly flows will also be examined.