Geophysical Research Abstracts
Vol. 19, EGU2017-3068-2, 2017
EGU General Assembly 2017

Comparison between stochastic and machine learning methods for hydrological multi-step ahead forecasting: All forecasts are wrong!

Georgia Papacharalampous (1), Hristos Tyralis (2), and Demetris Koutsoyiannis (3)

(1) National Technical University of Athens, School of Civil Engineering, Zografou, Greece (papacharalampous.georgia@gmail.com), (2) National Technical University of Athens, School of Civil Engineering, Zografou, Greece (montchrister@gmail.com), (3) National Technical University of Athens, School of Civil Engineering, Zografou, Greece (dk@itia.ntua.gr)

Machine learning (ML) is considered to be a promising approach to hydrological processes forecasting. We conduct a comparison between several stochastic and ML point estimation methods by performing large-scale computational experiments based on simulations. The purpose is to provide generalized results, while the respective comparisons in the literature are usually based on case studies. The stochastic methods used include simple methods, models from the frequently used families of Autoregressive Moving Average (ARMA), Autoregressive Fractionally Integrated Moving Average (ARFIMA) and Exponential Smoothing models. The ML methods used are Random Forests (RF), Support Vector Machines (SVM) and Neural Networks (NN). The comparison refers to the multi-step ahead forecasting properties of the methods.

A total of 20 methods are used, among which 9 are the ML methods. 12 simulation experiments are performed, while each of them uses 2 000 simulated time series of 310 observations. The time series are simulated using stochastic processes from the families of ARMA and ARFIMA models. Each time series is split into a fitting (first 300 observations) and a testing set (last 10 observations). The comparative assessment of the methods is based on 18 metrics, that quantify the methods' performance according to several criteria related to the accurate forecasting of the testing set, the capturing of its variation and the correlation between the testing and forecasted values.

The most important outcome of this study is that there is not a uniformly better or worse method. However, there are methods that are regularly better or worse than others with respect to specific metrics. It appears that, although a general ranking of the methods is not possible, their classification based on their similar or contrasting performance in the various metrics is possible to some extent. Another important conclusion is that more sophisticated methods do not necessarily provide better forecasts compared to simpler methods. It is pointed out that the ML methods do not differ dramatically from the stochastic methods, while it is interesting that the NN, RF and SVM algorithms used in this study offer potentially very good performance in terms of accuracy. It should be noted that, although this study focuses on hydrological processes, the results are of general scientific interest.

Another important point in this study is the use of several methods and metrics. Using fewer methods and fewer metrics would have led to a very different overall picture, particularly if those fewer metrics corresponded to fewer criteria. For this reason, we consider that the proposed methodology is appropriate for the evaluation of forecasting methods.