Investigation of the effect of the hyperparameter optimization and the time lag selection in time series forecasting using machine learning algorithms

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)

The hyperparameter optimization and the time lag selection are considered to be of great importance in time series forecasting using machine learning (ML) algorithms. To investigate their effect on the ML forecasting performance we conduct several large-scale simulation experiments. Within each of the latter we compare 12 methods on 2 000 simulated time series from the family of Autoregressive Fractionally Integrated Moving Average (ARFIMA) models. The methods are defined by the set {ML algorithm, hyperparameter selection procedure, time lags}. We compare three ML algorithms, i.e. Neural Networks (NN), Random Forests (RF) and Support Vector Machines (SVM), two procedures for hyperparameter selection i.e. predefined hyperparameters or defined after optimization and two regression matrices (using time lag 1 or 1, ..., 21). After splitting each simulated time series into a fitting and a testing set, we fit the models to the former set and compare their performance on the latter one. We quantify the methods' performance using several metrics proposed in the literature and benchmark methods. Furthermore, we conduct a sensitivity analysis on the length of the fitting set to examine how it affects the robustness of our results.

The findings indicate that the hyperparameter optimization mostly has a small effect on the forecasting performance. This is particularly important, because the hyperparameter optimization is computationally intensive. On the other hand, the time lag selection seems to mostly significantly affect the methods performance when using the NN algorithm, while we observe a similar behaviour for the RF algorithm albeit to a smaller extent.