# Prediction Of Hydrological Models' Uncertainty By A Committee Of Machine Learning-Models

Kayastha, Nagendra; Solomatine, Dmitri; Lal Shrestha, D.

**Publication date**
2014

**Document Version**
Final published version

**Published in**
Proceedings of the HIC 2014 - 11th international conference on hydroinformatics

8-1-2014

# Prediction Of Hydrological Models' Uncertainty By A Committee Of Machine Learning-Models

Nagendra Kayastha

Dimitri P. Solomatine

Durga Lal Shrestha

11th International Conference on Hydroinformatics
HIC 2014, New York City, USA

# PREDICTION OF A HYDROLOGICAL MODEL'S UNCERTAINTY BY A COMMITTEE OF MACHINE LEARNING-MODELS

NAGENDRA KAYASTHA (1), DIMITRI SOLOMATINE (1, 2), DURGA LAL SHRESTHA (3),

*(1): UNESCO-IHE Institute for Water Education, 26211AX, Delft, Netherlands*

*(2): Water Resources Section, Delft University of Technology, Netherlands*

*(3): CSIRO Land and Water, Highett, Australia*

In the MLUE method (reported in Shrestha *et al*. [1, 2]) we run a hydrological model *M* for multiple realizations of parameters vectors (Monte Carlo simulations), and use this data to build a machine learning model *V* to predict uncertainty (quantiles) of the model *M* output. In this paper, for model *V*, we employ three machine learning techniques, namely, artificial neural networks, model tree, locally weighted regression which leads to several models results. We propose to use the simple averaging method (SA) and the weighted model averaging method (WMA) to form a committee of these models. These approaches are applied to estimate uncertainty of streamflows simulation in Bagmati catchment in Nepal. Tests on the different data sets show that WMA performs a bit better than SA.

**Keywords**: uncertainty analysis, hydrological model, machine learning, MLUE, model averaging.

## INTRODUCTION

The concept of multi-model averaging is applied to combine machine learning models for uncertainty prediction of hydrological models built by the MLUE method (Shrestha et al., 2009). The basic idea of combining different predictive uncertainty models is to use the available information efficiently and to construct an averaged predictive uncertainty model with the right balance between model flexibility and overfitting.

Multi-model averaging is receiving attention in the hydrological modelling explicitly to derive predictive model output. The motivation behind multi-model averaging is to extract as much information as possible from existing competing models to produce a better output. Analysis of results from group of competing models is much more complex than any single model. Each model having its own predictive capabilities and limitations therefore it is difficult to compare. However, combination of competing models allows the strength of each individual model merging in optimal way so that it can obtain best prediction. Combining models require weights which averages the model outputs taking advantages of each individual model.

Uncertainty analysis of hydrological models mostly focuses on sampling based method where ensemble of deterministic model outputs generate to characterizes and quantifying the uncertainty. Machine learning techniques have been used to encapsulate results of MC simulations by building a predictive uncertainty models. The machine learning based uncertainty prediction approach is very useful for estimation of hydrological models' uncertainty in particular hydro-metrological situation in real-time application. In this approach, the hydrological model realizations from Monte Carlo simulations are used to build different machine learning uncertainty models to predict uncertainty (quantiles of pdf) of the a deterministic output from hydrological model. Uncertainty models are trained using antecedent precipitation and streamflows as inputs. The trained models are then employed to predict the model output uncertainty that is specific for the new input data.

This approach can be used results of any sampling scheme to build a machine learning model and able to predict uncertainty of a hydrological model outputs. The trained model called a predictive uncertainty model (*V*) that maps the input data to the prediction interval of the model output that is generated by sampling schemes. The details of methodology can be found in Shestha *et al.* [1, 2].

In this study, we present results of hydrological model outputs uncertainties predicted from number of machine learning models. Three machine learning models, namely artificial neural networks, model tree, locally weighted regression (ANN, MT, LWR) with six different model inputs structure are tested to predict uncertainty of streamflows simulation from a conceptual hydrological model HBV for Bagmati catchment in Nepal. The problem here is that several input datasets used to train model *V* (resulting in several models, total 18 models) and these are difficult to compare. We propose to form a committee of all predictive uncertainty models using averaging schemes to generate the single (final) output. Two schemes simple averaging (SA) and weighted model averaging (WMA, e. g., Ajami *et al*. [3] Shamseldin *et al*. [4]) methods are used in this study.

**Uncertainty prediction models and their averaging**

Model *V* encapsulating the functional relationship between the inputs and the prediction interval *PI* taken as following form:

$$PI^{k} = V_{k}(X_{u}) + \xi = \hat{PI}^{k} + \xi \quad (1)$$

where $PI^k$ is the prediction interval computed from MC data; $X_u$ is input for uncertainty prediction models $\hat{PI}^{k}$ is the prediction interval estimated by machine learning techniques; $k \in \{L, U\}$; *L*-lower and *U*-upper; $\xi$ is the residual error in estimating the prediction intervals.

WMA is technique to combine multiple models for better prediction among various competing models. The main idea of WMA is that the ensemble outputs generated by various models combined based on their performance. The WMA for combining multiple models of prediction intervals expressed as:

$$PI_{wma}^{k} = \sum_{n=1}^{N} w_n V_n^k \quad (2)$$

where *n* is individual uncertainty prediction model (*n*= 1..*N*), N is the number of models under consideration, *k* is *L*-lower and *U*-upper, $w_n$ is WMA weight better performing predictions receive higher weights than the worse performing ones. All weights are positive and should add up to 1. In SA method, the multiple models of prediction intervals obtained through simply arithmetic mean of considered models that is each of the models is weighted equally.

**Results and discussion**

Based on the average mutual information and correlation analysis several structures of input data sets considered for the machine learning models. Various combinations of the three effective rainfall values ($RE_{t-0}$ , $RE_{t-1}$ and $RE_{t-2}$,) and past values of the observed discharges ($Q_{t-1}$ and $\Delta Q_{t-1}$.) are considered as inputs. Table 1 presents six possible combinations of input structure used for the each machine learning model. It produces all in total 18 uncertainty prediction models.

Table.1. Input data structures of machine learning models to reproduce MCS uncertainty results of the HBV model

| Models | Input combination for uncertainty prediction models ($X_u$) |
|---|---|
| V01 | $RE_{t-0}$, $Q_{t-1}$ |
| V02 | $RE_{t-0}$, $Q_{t-1}$, $\Delta Q_{t-2}$ |
| V03 | $RE_{t-0}$, $RE_{t-1}$, $Q_{t-1}$, $Q_{t-2}$ |
| V04 | $RE_{t-0}$, $RE_{t-1}$, $Q_{t-1}$, $\Delta Q_{t-1}$ |
| V05 | $RE_{t-0}$, $RE_{t-1}$, $RE_{t-2}$, $Q_{t-1}$, $Q_{t-2}$ |
| V06 | $RE_{t-0}$, $RE_{t-1}$, $RE_{t-2}$, $Q_{t-1}$, $\Delta Q_{t-1}$ |

WMA is applied for combining 18 individual predictive uncertainty models based on six different input structures with three machine learning models (ANN, MT and LWR) for calibration and validation periods which are tested in Bagmati catchment. These results are presented in Table 2. The outputs generated by various models are combined using WMA. Each predictive model (e,.g, for lower PI) receives a weights which are calculated based on CoC (Coefficient of correlation). The averaging models are evaluated based on the prediction interval coverage probability (PICP) (should be close to the prescribed degree of confidence) and the mean prediction interval (MPI) (If there is no uncertainty, then *MPI* is zero).

.

$$PICP = \frac{1}{n}\sum_{t=1}^{n} C \text{ ; where } C = \begin{cases} 1, & PL_t^L \le y_t \le PL_t^U \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

$$MPI = \frac{1}{n}\sum_{t=1}^{n}(PL_t^U - PL_t^L) \quad (4)$$

The result of WMA model shows that PICP is better if compared to MT and LWR models but the ANN model is between the best and the worst. The best of all models is ANN V05 that has CoC value of 0.89 and 0.87 in for calibration and verification respectively in lower PI and value of 0.96 and 0.95 respectively in upper PI, and the value of PICP is 74.43% in calibration and 78.77% in verification, which are highest among the all models. ANN V01 model received lower performance of CoC and PICP among all models in calibration and verification, however MPI is narrow (considered better than other models)..The WMA produced PICP 64.35 and 69.74 % in calibration and validation period respectively. However, it produced wider MPI among all models except ANN V05.

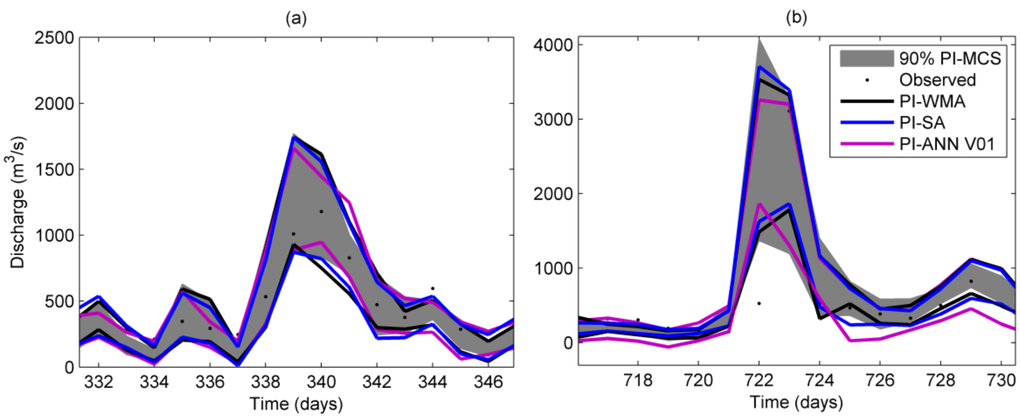


Figure 1. Hydrograph of 90% prediction bounds in verification period, the black dot indicates observed discharges and the dark grey shaded area denotes the prediction uncertainty that results from MCS. Black, blue and purple lines denote the prediction uncertainty estimated by WAM, SA and ANN-V01 respectively.

Table 2. Performances of the models predicting the 5 and 95% quantiles (Lower and higher PI respectively) in verification.

| ML techniques | Models | CoC | | RMSE | | PICP | MPI |
|---|---|---|---|---|---|---|---|
| | | Lower PI | Upper PI | Lower PI | Upper PI | | |
| ANN | V01 | *0.71* | *0.86* | 60.25 | 88.05 | *56.84* | *118.73* |
| | V02 | *0.71* | *0.86* | **60.79** | **92.09** | 75.52 | 142.80 |
| | V03 | 0.81 | 0.94 | 51.46 | 61.59 | 66.24 | 124.03 |
| | V04 | 0.81 | 0.94 | 49.96 | 60.81 | 68.91 | 125.79 |
| | V05 | 0.87 | 0.95 | *43.34* | 67.53 | **78.77** | **160.48** |
| | V06 | 0.82 | 0.93 | 49.54 | 66.28 | 73.32 | 136.94 |
| MT | V01 | 0.72 | 0.90 | 59.14 | 76.92 | 64.04 | 118.95 |
| | V02 | 0.73 | 0.90 | 58.68 | 76.81 | 66.24 | 119.14 |
| | V03 | 0.77 | 0.95 | 54.93 | 53.14 | 59.40 | 120.42 |
| | V04 | 0.76 | 0.95 | 55.66 | 53.27 | 60.09 | 119.67 |
| | V05 | 0.81 | 0.95 | 50.25 | 52.14 | 59.05 | 120.59 |
| | V06 | 0.80 | 0.95 | 51.18 | 52.21 | 59.51 | 119.89 |
| LWR | V01 | *0.71* | 0.89 | 59.80 | 78.42 | 61.37 | 120.19 |
| | V02 | 0.74 | 0.90 | 57.12 | 73.83 | 58.82 | 118.65 |
| | V03 | **0.86** | **0.96** | 44.56 | 50.37 | 59.16 | 121.73 |
| | V04 | **0.86** | **0.96** | 44.42 | 51.09 | 57.89 | 121.01 |
| | V05 | 0.87 | **0.96** | 43.33 | 49.62 | 59.74 | 123.05 |
| | V06 | **0.86** | **0.96** | 44.10 | 49.85 | 59.28 | 122.33 |
| SA | | 0.79 | 0.93 | 52.14 | 64.11 | 63.57 | 125.24 |
| WMA | | **0.86** | 0.94 | 45.79 | *47.84* | 69.74 | 136.20 |

## Conclusion

We are building predictive uncertainty models *V* to encapsulate the relationship between the hydrometeorological variables and the quartiles of the model output probability distribution (forming the prediction interval). MC sampling for uncertainty estimations are done off-line only to generate the data to train the model *V*, while the trained *V* models are employed to estimate the uncertainty in real time application without running the any sampling based simulations any more.

It is not straightforward to compare the results of many predictive uncertainty models. WMA overcomes the problem by conditioning, not on single best model but on the entire group of models. We show one of the ideas of model averaging which can be employed to combine several predictive uncertainty models. WMA for combining different predictive uncertainty models leads to increase in accuracy. It is observed that the percentage of the observation discharge data falling within the prediction bounds is highest for WMA. The verification results show that both averaging methods in general improve the predictive performance, but WMA is a bit better than SA.

## Acknowledgments

Part of this work was financially supported by the WeSenseIt and MyWATER projects supported by the FP7 Programme of the European Commission.

## REFERENCES

[1] Shrestha D. L., Kayastha N., Solomatine D. P., "A novel approach to parameter uncertainty analysis of hydrological models using neural networks", *Hydrology and Earth System Sciences,* 13, (2009), 1235–1248.

[2] Shrestha D. L., Kayastha N., Solomatine D. P. and Price. R., "Encapsulation of parameteric uncertainty statistics by various predictive machine learning models: MLUE method", *Journal of Hydroinformatic*, 16, 1, (2013). 95–113.

[3] Ajami N. K., Duan Q., Gao X., Sorooshian S.,: "Multimodel Combination Techniques for Analysis of Hydrological Simulations: Application to Distributed Model Intercomparison Project Results", *Journal of Hydrometeorology*, **7**, (2006),755–768.

[4] Shamseldin A. Y., O'Connor K. and Liang G. C., "Methods for combining the outputs of different rainfall–runoff models", *Journal of Hydrology*, 197**,** (1997) 203–229.