Review of HESS-2014-448: "Improving inflow forecasting into hydropower reservoirs through a complementary modelling framework"

This paper presented a complementary modelling framework that aims to improve accuracy of hourly reservoir inflow forecasts of deterministic hydrological models. The approach estimates the uncertainty in the complementary model structure and produces probabilistic inflow forecasts. The proposed method is practically useful, however, theoretical improvement has not been found and some questions remain:

Major comments:

(1) The authors claim that the described complementary conceptual and data-driven (error) models is a new approach. However, as stated in Lines 4-5, Page 12067, "Several example applications can be found in the scientific literature on using conceptual and data driven models complementarily", similar works have been found in the previous studies. Furthermore, the HBV model for conceptual model and autoregressive (AR) model for error model are both very mature models in hydrology. Therefore it is hard to find the new contribution or improvement in this paper.

(2) Actually, there are many error models at present, e.g. autoregressive nodel, autoregressive threshold model, fuzzy autoregressive threshold model, ARIMA based error models and artificial neural network models, and so on. This paper selected the autoregressive model to describe the error processes. The reason or additional statement should be given to be clear to the readers. More error models should be used and compared to obtain more reasonable and high accuracy results.

(3) This paper attempted to produce probabilistic inflow forecasts through a complementary modelling framework. However, it is known to all that the Bayesian forecasting system (BFS) and generalized likelihood uncertantity estimation (GLUE) may be the two most popular and widely used frameworks to produce probabilistic inflow forecasts. Comparisons of the results of the proposed method and the two methods mentined above are neseesary to verify whether the proposed method are more effective and reliable or not?

Minor comments:

(1) As shown in Fig. 8, the unit of inflow should be tranformed to international unit " m^3/s ".

(2) Some indexes in the following references can help identify and evaluate the quality of prediction interval, such as the percentage of coverage (POC), the average relative width (ARW) etc.

Xiong, L.H., Wan, M., Wei, X.J., and O'Connor, K.M. (2009). "Indices for assessing the prediction bounds of hydrological models and application by generalised likelihood uncertainty estimation." Hydrol. Sci. J., 54 (5): 852-871.

Li, L., Xu, C. Y., and Engeland, K. (2013). "Development and comparison in uncertainty assessment based Bayesian modularization method in hydrological modeling." J. Hydrol., 486, 384-394.

(3) Relative error (*RE*) is suggested to be used in the conceptual model during the calibration and validation period (Table 2). The value of RE is expected to be close to zero for a good simulation of the total volume of the observed runoff series, defined as

$$RE = \frac{\sum (Q_t - \hat{Q}_t)}{\sum Q_t} \times 100\%$$

where Q_t is the observed discharge at time *t*, $\hat{Q}_t$ is the corresponding simulated discharge.