Acclimatizing Fast Orthogonal Search (FOS) Model for River Stream-flow Forecasting

We would like to thank the referee for his objective and thorough review of our paper. We have addressed all the referee's comments in the following point-by-point response. All changes made to accommodate the referee's comments are underlined in the revised manuscript.

Reviewer #1

The manuscript presents a new method namely Fast Orthogonal Search (FOS) for inflow forecasting, which is interesting. The model is tested using inflow data from the Aswan High Dam located in Egypt. The subject addressed is within the scope of the journal. Overall, I think the paper is well written and the authors address an important topic in hydrology (inflow forecasts) that is of keen interest of the Hydrology community

Based solely on the paper results, I am not fully convinced that the proposed FOS model has much advantage (if any) over the classical AR and ARMA (including periodic AR) even the Artificial Neural Network model classes of inflow forecast models. Some results are NOT addressed and discussed adequately. Moreover, the apparent relative medium forecasting skill of the proposed model needs to be discussed.

However the manuscript, in its present form has the potential for publication in HESS with adequate revisions to the following points which should be undertaken in order to justify recommendation for publication.

Reply

The author thanks the reviewer for his comments. The authors address all of his comments one-by-one hereafter and modify the manuscript

---

- For readers to quickly catch the contribution in this work, it would be better to highlight major difficulties and challenges, and your original achievements to overcome them, in a clearer way in abstract.

Reply

Owing to the referee feedback, challenges and difficulties about the flow forecasting and also the original achievements have been reported more clearer way in the introduction section.

- Many assumptions are stated in various sections. More justifications should be provided on these assumptions. Evaluation on how they will affect the results should be made.

Reply

It is true that there are some assumptions in our research. Hereafter, we will try to highlight the major ones.

- **Assume the training approach.**

The findings of the cross-correlation analysis for the monthly natural inflow pattern for consequences years shows that the cross-correlation is relatively poor if go more than one year behind the one under study to be forecasted for most of the months. Based on that observation, theoretically, for the forecasting model has been developed based on the previous year inflow pattern for all training approaches used as training period and the followed forecasting period.

- **Assumed performance indicators**

Actually, in developing such forecasting model utilizing time series concept, the model could perform well during the training period and might provide higher level of error when evaluating during either validation or testing period. In this context, in this study the authors used these performance indices to make sure of that the proposed model could provide consistent level of accuracy during all periods.

The advantages of utilizing these two statistical indices as a performance indicator of the proposed model are as follow:-

1- **Using the maximum error is to make sure that the highest error while evaluating the performance is within the acceptable error for such forecasting model.**
2- **While utilizing the Root Mean Square error is to ensure that the summation of the error distribution within the validation period is not high.**
3- **Consequently, using both indices is guaranteed consistent level of errors which is providing a great potential for having same level error while examining the model for unseen data in the testing period.**

- The key FOS parameters are not mentioned. The rationale on the choice of the particular set of parameters should be explained. Have the authors experimented with other sets of values? What are the sensitivities of these parameters on the results?

Reply

In fact, there is no formal and/or mathematical method for determining the appropriate "optimal set" set of the key parameters of FOS which are four Model Order, Maximum Delay, Mean Square Error, and Mean Square Error Reduction. Accordingly, the authors decide to perform this task utilizing trial and error method. The authors experimented several sets and examined each experiment but we report only the best trial.

However, the authors reported some observations about the proposed model performance and sensitivity analysis under different set of key parameters in the revised version of the manuscript. Details of four key parameters are given below.

"There are four main parameters governing the NARMAX model formed by FOS. The First two parameters are the model order and the maximum delay of the NRMAX model. The maximum model order in this research was initially set to 5 and the maximum model order obtained was 3. The maximum delay was set to 12 and was decided based on the data feeding process. The maximum delay of the obtained models did not exceed 6. The third

parameter is the minimum MSE obtained and this was set to 10^{-6} to ensure best fitting of data. The fourth parameter is the MSE reduction introduced by each term. This parameter value is internally determined by FOS."

- It is mentioned in p.4 line 17-19 that "Even though these AI models demonstrate to be proficient, the convergence of the model during the training (calibration) experienced a slow procedure which means that the model falls in the sub-optimal search procedure.". Some justifications should be furnished on this.

Reply

The authors fully agreed with the referee in this point that the statement is not fully understandable. In this context, the authors add more clarification in this position of the manuscript. In most of Artificial Neural Network (ANN) models development the back-propagation algorithm is used for optimizing the ANN key parameters. The back-propagation algorithm experienced several drawbacks such as, local optima, slowness. In fact, there are many advanced methods offered by researchers to overcome partially these drawbacks especially the local optima such as Particle Swarm Optimization (PSO) and Genetic Algorithm (GA). However, utilizing those proposed optimization algorithm to treat the drawback of the back-propagation experienced another challenges such as over-fitting problem for the whole model performance.

- There is missing information about the major feature of the Nile basin and proper statistical analysis for the data. It is vital for the readers to get complete information about the basin and also brief statistical analysis for the raw data.

Reply

Owing to the reviewer feedback, the authors add comprehensive description for the Nile River Basin has been reported in the case study section. In addition, statistical analysis for the natural inflow pattern for 130 years at AHD has been carried out.

- The authors presents the three different training approaches for the model and shows graphically its time-line procedure, however there is absence of the major mode structure for the model, even the authors describe the model structure satisfactorily, it would be better to show a block diagram for the model structure.

Reply

Block diagram for the model structure has been added

- It would be of importance for the readers to see more performance indicators for the model evaluation to be presented. In addition, as long as the authors are presenting the results on monthly basis, it would be better to do so for each month.

Reply

Owing to the reviewer feedback, the authors add one more table " table " to show the complete performance for the proposed FOS model showing the performance indicators (four performance indicators" for each month.

- Moreover, the manuscript could be substantially improved by relying and citing more on recent literatures about case studies of application of various types of soft computing technique in discharge prediction such as the followings:

- Cheng, C.T., Wu, X.Y. and Chau, K.W., "Multiple criteria rainfall-runoff model calibration using a parallel genetic algorithm in a cluster of computer," Hydrological Sciences Journal, Vol. 50, No. 6, 2005, pp. 1069-1087.

- Lin, J.Y., Cheng, C.T. and Chau, K.W., "Using support vector machines for long-term discharge prediction," Hydrological Sciences Journal, Vol. 51, No. 4, 2006, pp. 599-612.

- Wang, W.C., Chau, K.W., Cheng, C.T. and Qiu, L., "A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series," Journal of Hydrology, Vol. 374, No. 3-4, 2009, pp 294-306.

- Wu, C.L., Chau, K.W. and Li, Y.S., "Predicting monthly streamflow using data-driven models coupled with data-preprocessing techniques," Water Resources Research, 45, W08432, doi:10.1029/2007WR006737, 2009.

- Cheng, C.T., Ou, C.P. and Chau, K.W., "Combining a fuzzy optimal model with a genetic algorithm to solve multiobjective rainfall-runoff model calibration," Journal of Hydrology, Vol. 268, No. 1-4, 2002, pp. 72-86.

- Chau, K.W., "Particle swarm optimization training algorithm for ANNs in stage prediction of Shing Mun River," Journal of Hydrology, Vol. 329, No. 3-4, 2006, pp. 363-367.

Reply

All the above references have been reviewed and included in the revised manuscript.

- Complete results for all the performance indicators should be presented in the discussion section.

Reply

The authors improve the results and discussion section adding more details discussion on the model performance.

- In the conclusion section, the limitations of this study, suggested improvements of this work and future directions should be highlighted.

Reply

The conclusion section has been improved and includes the limitations of this study, suggested improvements of this work and future directions.
