We thank the reviewers for their insightful comments and constructive suggestions that have led to the improvement of our paper. Our responses to all the comments and suggestions are detailed below.

Reviewer #1:

This manuscript proposed an AI framework consisting of hyperparameter selection and training parts to improve training efficiency and reduce overfitting. And the case study for daily runoff prediction in the Maumee domain showed the good potential of this framework. In general, this manuscript is well-written, and the conclusion is reasonable. Therefore, I think the manuscript can be published in EGUsphere after minor revision.

(a) When discussing the hyperparameter SS, the paper mentioned that the daily EOF runoff is imbalanced data. Usually, this is very important for AI model training. Can authors explain how to deal with the imbalanced data? I recommend the authors preprocess the data to improve the availability of the data in model training if the authors did not deal with the imbalanced data.

Thanks. We agreed with the reviewer that the imbalanced data poses great challenges to model training. There are several ways to improve the training quality for imbalanced data: 1) Select an effective machine learning (ML) algorithm that has built-in mechanisms to deal with imbalanced data. 2) Choose a good cross-validation strategy to ensure the training and test datasets follow a similar distribution of the target variable. 3) Set class weights on your target classes to give more weight to the minority class. In our study, as our focus is to identify the influential hyperparameters for the regression models that are trained to predict the magnitude of daily EOF runoff, we chose an effective ML algorithm, the Extreme Gradient Boosting (XGBoost) algorithm for model training; the XGBoost algorithm offers a range of hyperparameters that can give fine-grained control over the model training performance against imbalance data. For example, we used the Stratified K-Fold cross-validation to ensure the training and test datasets follow a similar distribution and defined a loss function that penalizes more the missing predictions of non-zero runoff events, that is, the minority class in this study.

(b) I suggest authors add a flowchart in the manuscript, which will be good for readers to understand the framework.

Thanks. Figure 1 explains the different components of the methodological framework. To further clarify, we follow the suggestion by the reviewer to add some descriptions on the workflow as follows:

We first choose a machine learning algorithm and its associated hyperparameters. Then, we feed the initial hyperparameters (1) to the hyperparameter selection (HS) module to determine the influential hyperparameters (2). Once initial values are assigned to the influential hyperparameters (4), we use the hyperparameter tuning to identify their optimal values (3), which allows the algorithm to achieve the best performance in training. A case study is used to illustrate the workflow in more detail.