Response to referee comments on "Modelling salinity in river systems using hybrid process and data-driven models", by Jason M. Hunter et al.

Note: This document contains the authors' responses to the comments of Referee #1. The comments made by the referee have been formatted in italic and coloured black, while our responses are upright and coloured blue.

Response to the comments of Referee #1

This study developed hybrid process and data-driven model to improve single-driven model performance for modelling salinity in river systems. Despite the paper is well organized and interesting to read, the manuscript in its present form has some weaknesses (mainly lack novelty and scientific findings).

We would like to thank the reviewer for their constructive comments, which will assist with improving the quality of the paper significantly. Detailed responses to the reviewer's comments are given below.

(1) The introduction and methodology (13 pages) are too long. Please make it concise and shorter and emphasize the novelty of the study.

We agree that the novelty of the study could be articulated more clearly. This will be done in the revised version by:

1. Changing the title of the paper to "Framework for developing hybrid process and data driven (artificial neural network and regression) models of salinity in river systems", thereby highlighting that the primary contribution of the paper is the framework introduced in Section 2, the application of which is illustrated for a real case study in the River Murray, Australia.
2. The fact that the development of the generic framework is the primary contribution of the paper will be highlighted in the revised version of the Abstract.
3. The objectives of the paper, highlighting the contribution of the framework, will be stated explicitly in the Introduction.
4. We will review sections 1 to 3 carefully and make every effort to make them as clear and concise as possible. For example, the paragraph on lines 72 – 81 will be deleted from the Introduction.

The proposed framework (Section 2) and the demonstration of how this is applied to a real case study (Section 3) are the primary contributions of the paper, so most of the paper is devoted to these topics. We believe that this is appropriate and will be made clear by the proposed changes to the Title, Abstract and Introduction outlined above.

(2) The results and discussion (1 page excluding tables and figures) are too shorter. Please enrich it and offer more valuable analyses and scientific findings.

As mentioned in our response to Comment (1), the framework and a demonstration of how this is applied to a real case study are the primary contributions of the paper. In contrast, the purpose of the actual modelling results is to demonstrate the utility of the proposed approach (or otherwise) and is hence quite brief by design. However, additional discussion in relation to the advantages and disadvantages of the proposed framework, which is the primary contribution of the paper, will be provided in the Results and Discussion section in the revised version of the paper.

(3) In Figure 5, the descriptions of “Below 30,000” A ¸A“Below 50,000”etc, are imprecise. Please replace it with ‘Below 30,000 and above 15,000’ etc. Besides, the symbol “sigma” easily causes readers’ misunderstanding that the results of Model 1 are equal to the sum of Model 2-5. Please make major revision for Figure 5.

The clarity of Figure 5 will be improved in the revised version of the paper in accordance with the reviewer’s suggestions.

(4) In Eq. (3), so many researchers suggested that it needs use the index Gbench (or Coefficient of Persistence) by replacement of NSE to judge the good-of-fit of the model, when you applied a data-driven model, such as ANN on the basis of benchmark series. Please refer to the reference “Seibert, J. (2001). On the need for benchmarks in hydrological modelling. Hydrological Processes, 15(6), 1063-1064”.

The goodness-of-fit statistic suggested by the reviewer will be added in the revised version of the paper.

(5) In Figure 8, the inputs are so important for data-driven model. Why you ANN model just has the exogenous inputs (ex, Lock 5: electrical conductivity with 5-day lag, Lyrup pump station: water level with 3-day lag, Lock 5: flow rate 5-day lag, and Lock 5: water level with 5-day lag), but hasn’t the autoregressive input (Lock 4: salinity with 1-day lag). As known, the contributions of autoregressive input for model performance are higher than 80%-90%, however, the contributions of the exogenous inputs for model performance are only 10%-20%. Please explain it.

This is a very important point and we would like to thank the reviewer for raising it. Whether autoregressive inputs are considered as candidate inputs or not is a function of the purpose of the model. If the purpose is to obtain the best possible forecasts, then autoregressive inputs should be included as candidate inputs, as suggested by the reviewer. However, if the purpose of the model is to predict an independent variable as a function of other variables, as is the case in the case study considered as the model is supposed to be used to assess the impact of different management options on salinity, then autoregressive inputs cannot be considered.

In relation to the proposed general framework, this point will be added to the discussion on “Model Purpose” in Section 2 (issues like this is the reason for the inclusion of the consideration of model purposes as part of the proposed framework).

In relation to the case study, the reason for not considering autoregressive candidate inputs will be explained in Section 3 in the revised version of the paper.

(6) In Figure 8, how do you identify the time-lags of inputs? Please add your methods and results to demonstrate their suitability.

We agree that this could be articulated more clearly. In Section 3.6.1, we state that the ANN benchmark model was developed using the same methodology as was used for the development of Model 2, referring to Section 3.5.4 (to ensure the results of the different models can be compared in an objective fashion). In section 3.5.4, we state that the relevant inputs are determined with the aid of correlation analysis. Consequently, the time lags of the inputs in Figure 8 were determined using correlation analysis. However, we will provide model development details in supplementary material for additional clarity and completeness in the revised version of the paper. Given the length of the paper and the primary focus on the proposed hybrid approach, rather than the development of the component models, for which well-developed methodologies already exist, we believe this is more appropriate than giving these details in the paper (which is the reason they were omitted from the first submission).

(7) Section 2.3, the methodology for identification of most suitable model types is not scientific and imprecise. From the results of Table 2, the most suitable model types are identified based on the degree of data availability and process understanding. How do you quantify the degree of data availability and process understanding? Please make major revision of section 2.3 for enhancing the reliability of this method.

We agree with the reviewer that ideally, there would be hard and fast rules to assist model developers in determining which model is most appropriate given their circumstances. However, given that we are proposing a generic framework that is designed to be applicable under a wide range of circumstances, this is not possible. The purpose of the proposed framework is to raise these issues as steps that modellers must follow. However, inevitably, a degree of judgement will be required, given the high degree of variability in modelling contexts. In this sense, the concepts introduced are similar to the well-known figure of Grayson and Blöschl (2000) referred to in the paper and shown here, where it is not possible to provide precise quantitative values.

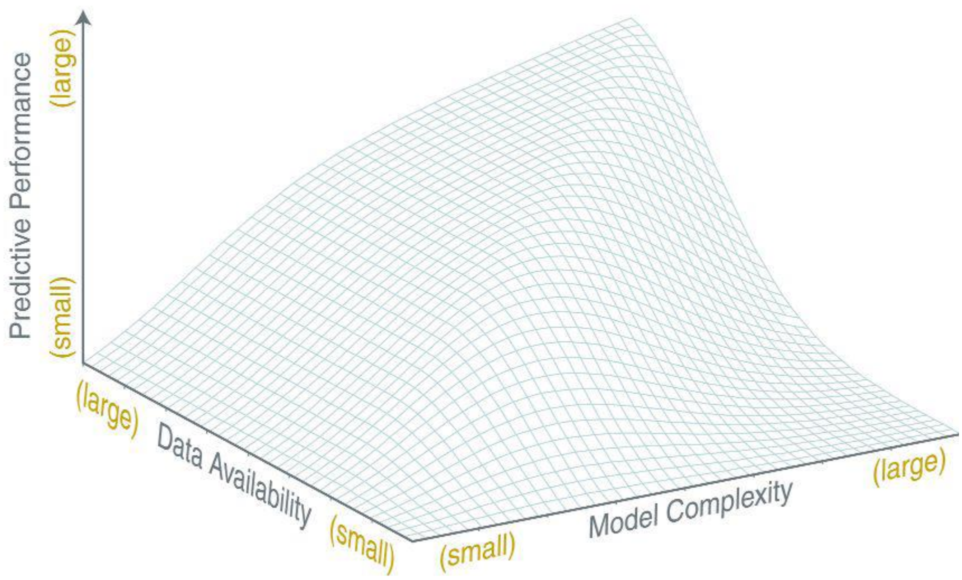

Figure 1: Relationship Between Data Availability, Model Complexity and Predictive Performance (Grayson and Blöschl, 2000).

This will be made clearer in the revised version of the paper in a number of places, namely Sections 2.1, 2.3 and 4.

Reference:

Grayson, R. B., and Blöschl, G.: Spatial Patterns in Catchment Hydrology: Observations and Modelling, Cambridge University Press, Cambridge, United Kingdom, 2000.

(8) In Page 20, lines 438-440: Please add the results of trials with one to four hidden nodes to demonstrate that ANN with three hidden nodes preforms best on the calibration data. Providing the results of RMSE and NSE.

These results will be provided as supplementary material in the revised version of the paper. As mentioned in our response to Comment (6), we believe this to be most appropriate, given the length and focus of the paper.

(9) You stated the limitation of your methodology "While the approach has been developed specifically for the modelling of salinity in rivers, ...". In fact, this methodology just developed specially for modelling salinity in Murray River of South Australia. Hence, the title of this paper might be changed as follow. Modelling salinity in Murray River of South Australia using hybrid process and data-driven models.

We agree that the purpose and contribution of the paper was not articulated as clearly as it should have been. However, as per our responses to Comments (1) and (2), the paper introduces a generic framework that is illustrated using the River Murray case study. This will be made clear by the changed title and clearly stated objectives (see responses to Comment (1)).