Developing and Testing a Long Short-Term Memory Stream Temperature Model in Daily and Continental Scale

Farshid Rahmani[1], Samantha Oliver[2], Wenyu Ouyang[3], Alison Appling[4], Kathryn Lawson[1], and Chaopeng Shen[1]

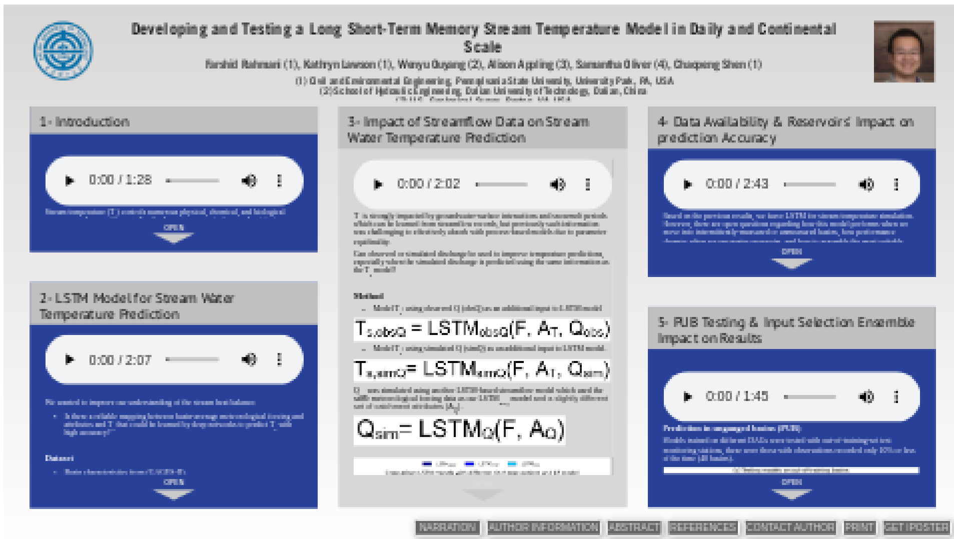
[1]Pennsylvania State University Main Campus
[2]USGS
[3]DUT Dalian University of Technology
[4]USGS Integrated Information Dissemination Division

November 21, 2022

Abstract

Stream water temperature (T) is a variable of critical importance and decision-making relevance to aquatic ecosystems, energy production, and human's interaction with the river system. Here, we propose a basin-centric stream water temperature model based on the long short-term memory (LSTM) model trained over hundreds of basins over continental United States, providing a first continental-scale benchmark on this problem. This model was fed by atmospheric forcing data, static catchment attributes and optionally observed or simulated discharge data. The model achieved a high performance, delivering a high median root-mean-squared-error (RMSE) for the groups with extensive, intermediate and scarce temperature measurements, respectively. The median Nash Sutcliffe model efficiency coefficients were above 0.97 for all groups and above 0.91 after air temperature was subtracted, showing the model to capture most of the temporal dynamics. Reservoirs have a substantial impact on the pattern of water temperature and negative influence the model performance. The median RMSE was 0.69 and 0.99 for sites without major dams and with major dams, respectively, in groups with data availability larger than 90%. Additional experiments showed that observed or simulated streamflow data is useful as an input for basins without major dams but may increase prediction bias otherwise. Our results suggest a strong mapping exists between basin-averaged forcings variables and attributes and water temperature, but local measurements can strongly improve the model. This work provides the first benchmark and significant insights for future effort. However, challenges remain for basins with large dams which can be targeted in the future when more information of withdrawal timing and water ponding time were accessible.

Developing and Testing a Long Short-Term Memory Stream Temperature Model in Daily and Continental Scale

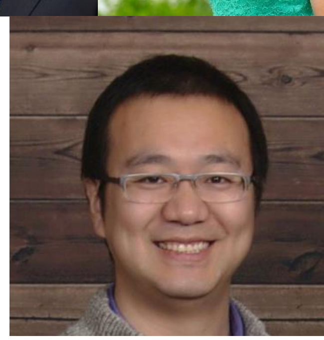


Farshid Rahmani (1), Kathryn Lawson (1), Wenyu Ouyang (2), Alison Appling (3), Samantha Oliver (4), Chaopeng Shen (1)

(1) Civil and Environmental Engineering, Pennsylvania State University, University Park, PA, USA
(2) School of Hydraulic Engineering, Dalian University of Technology, Dalian, China
(3) U.S. Geological Survey, Reston, VA, USA
(4) U.S. Geological Survey, Upper Midwest Water Science Center, Middleton, WI, USA







PRESENTED AT: AGU 2020

AGU FALL MEETING
Online Everywhere | 1–17 December 2020

1- INTRODUCTION

Stream temperature (T_s) controls numerous physical, chemical, and biological processes and properties, e.g. dissolved oxygen concentrations and nutrient transformation rates, as well as industrial processes such as cooling power plants and treating drinking water (Delpla et al. 2009; Madden, Lewis, and Davis 2013; Kaushal et al. 2010). Thermal regimes of streams directly affect aquatic species (Justice et al. 2017) and in some cases, fish mortality rate increases as water temperature passes a certain threshold for even a few hours (Martins et al. 2012; Marcogliese 2001). These are complicated by water uses in industry, such as utilizing stream water for cooling systems purposes, which causes thermal pollution downstream. Fulfilling the temperature requirements of the environment, agriculture, industries, and municipalities, and coordinating these uses requires a delicate balance. Accurate water temperature models can inform the decision making process and help lower the risks of exceeding the thermal thresholds, but there are many factors influencing water temperature, making modeling very complicated.

Long short-term memory (LSTM) is a deep learning algorithm, which has received increasing attention in hydrologic literature for its ability to learn and keep information for long periods. Based on LSTM, we developed multiple basin-centric lumped $T_s$models which were trained and tested on basins in the conterminous United States (CONUS)

In this poster, we attempt to briefly answer the following research questions:

(1) Is there a reliable mapping between basin-average meteorological forcing and attributes and Ts that could be learned by deep networks to predict Ts with high accuracy?

(2) Can observed or simulated discharge be used to improve temperature predictions, especially when the simulated discharge is predicted using the same information as the Ts model?

(3) How to improve the prediction results in ungauged basins?

(4) How much do reservoirs impact water temperature modeling?

2- LSTM MODEL FOR STREAM WATER TEMPERATURE PREDICTION

We wanted to improve our understanding of the stream heat balance:

- Is there a reliable mapping between basin-average meteorological forcing and attributes and T_s that could be learned by deep networks to predict T_s with high accuracy?

Dataset

- Basin characteristics from (GAGES-II).
- Basin daily meteorological forcing data (precipitation, maximum and minimum air temperature, vapor pressure, solar radiation) from Google Earth Engine (GEE).
- T_s observations from USGS national Water Information System (NWIS).
- 118 basins: only gauges with more than 60% of daily observations, in basins where there were no major dams.

Evaluation Metrics

- Loss function (model's goal was to minimize) was root-mean-square error (RMSE)
- Also report bias, correlation, and Nash-Sutcliffe efficiency coefficient (NSE)
- NSE_{res} and $Corr_{res}$ calculated by a new value to remove seasonal trends:

$T_{res} = T_s - T_{air}$

Method

- Model T_s, using basin-average climate forcings and attributes, without any data from streamflow (Q):

$T_{s,noQ}$ = $LSTM_{noQ}(F, A_T)$

Characteristics

- Four years training, two years testing
- Baseline for caomparison: a locally-fitted autoregressive model with exogenous variables (ARX_2):

$$T_s^{t,*} = \sum_{i=1}^{2} a_i T_s^{t-i,*} + \sum_{i=0}^{2} \sum_{j=1}^{p} b_{i,j} X_j^{t-i} + c$$

- X: atmospheric forcings (maximum and minimum air temperature, observed Q)
- a, b and c were fitted coefficients
- $T_s^{t-i,*}$ is the stream temperature simulated by ARX_2 at time step t
- p is the number of forcings

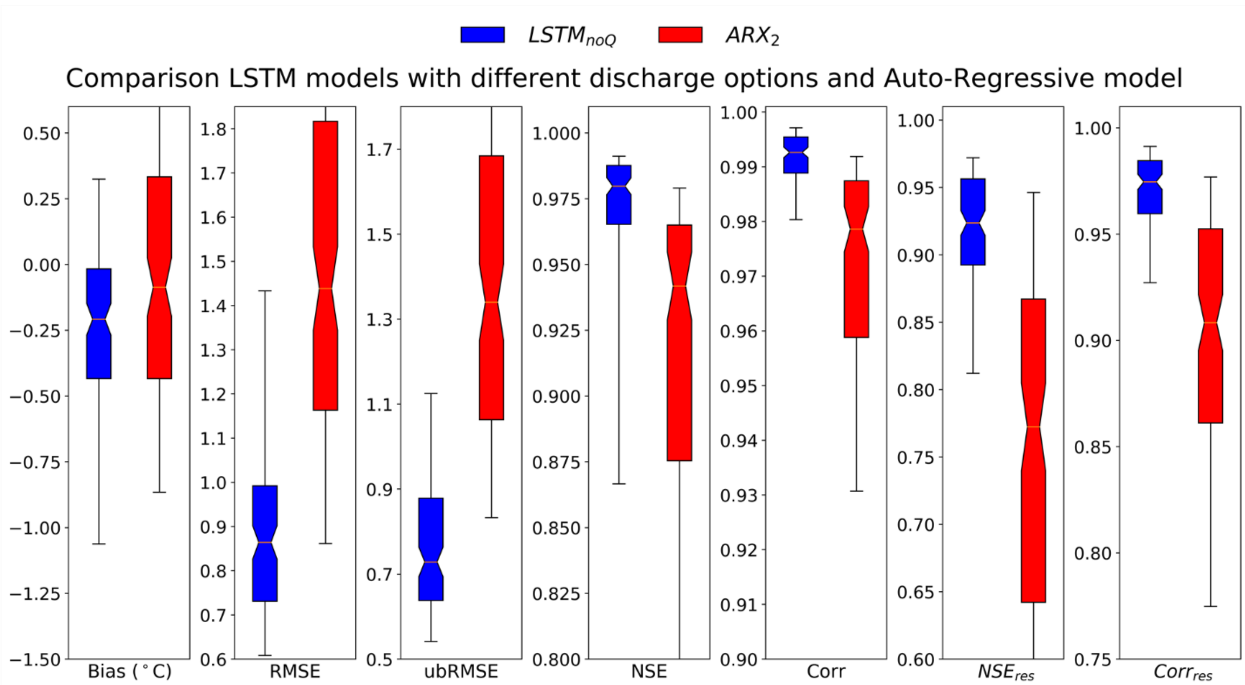


Figure 1. CONUS-scale aggregated metrics of stream temperature models for the test period. $LSTM_{noQ}$ had no input streamflow information, ARX_2 is the locally-fitted auto-regressive model.

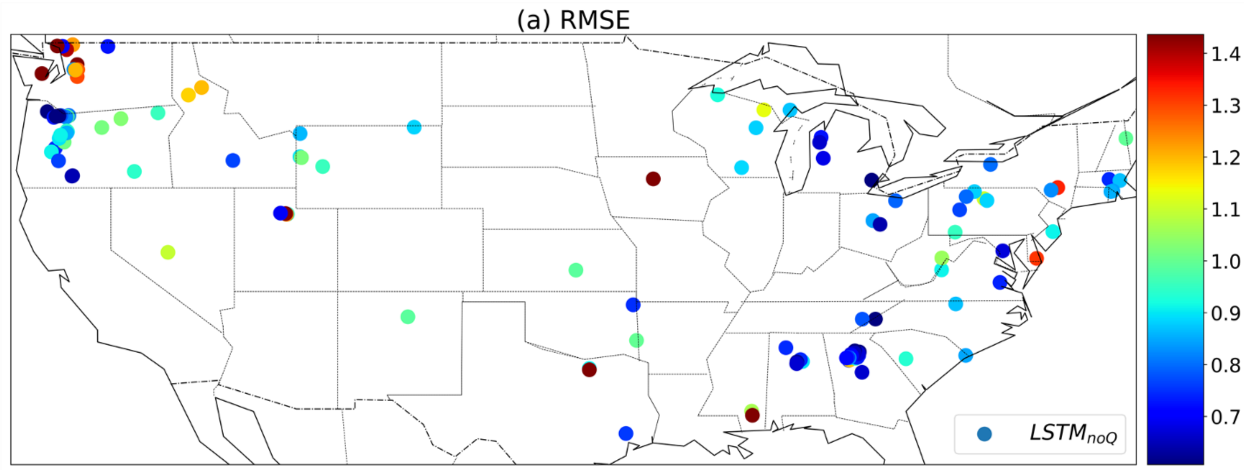


Figure 2. Spatial Distribution of RMSE

Result

- Exceptionally strong performance LSTM-based models (The median RMSE: 0.86 °C, and the correlation: 0.992).
- These metrics are markedly better than in previous studies
- Temporal fluctuations were extremely well captured (median NSE: 0.979).
- at this scale, which suggests that LSTM is particularly well-suited for Ts modeling at basin outlets.
- NSE_{res} improved substantially from ARX_2 (0.772) to LSTM (0.91) indicating that LSTM model was much less reliant on air temperature.
-

3- IMPACT OF STREAMFLOW DATA ON STREAM WATER TEMPERATURE PREDICTION

T_s is strongly impacted by groundwater-surface interactions and snowmelt periods which can be learned from streamflow records, but previously such information was challenging to effectively absorb with process-based models due to parameter equifinality.

Can observed or simulated discharge be used to improve temperature predictions, especially when the simulated discharge is predicted using the same information as the T_s model?

Method

- Model T_s: using observed Q (obsQ) as an additional input to LSTM model

$$T_{s,obsQ} = LSTM_{obsQ}(F, A_T, Q_{obs})$$

- Model T_s: using simulated Q (simQ) as an additional input to LSTM model.

$$T_{s,simQ} = LSTM_{simQ}(F, A_T, Q_{sim})$$

Q_{sim} was simulated using another LSTM-based streamflow model which used the same meteorological forcing data as our $LSTM_{noQ}$ model and a slightly different set of catchment attributes (A_Q).

$$Q_{sim} = LSTM_Q(F, A_Q)$$

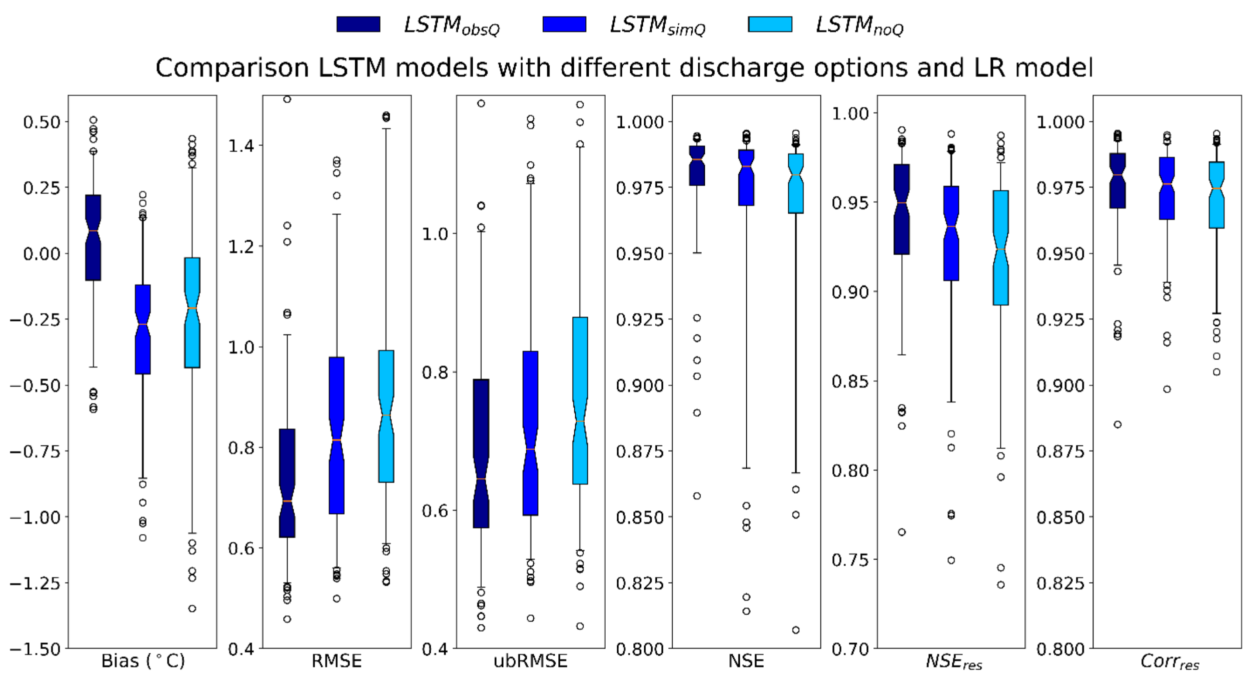


Figure 3. CONUS-scale aggregated metrics of stream temperature models for the test period. $LSTM_{obsQ}$ incorporated observed streamflow, $LSTM_{noQ}$ had no input streamflow information, while $LSTM_{simQ}$ incorporated simulated streamflow (simQ).

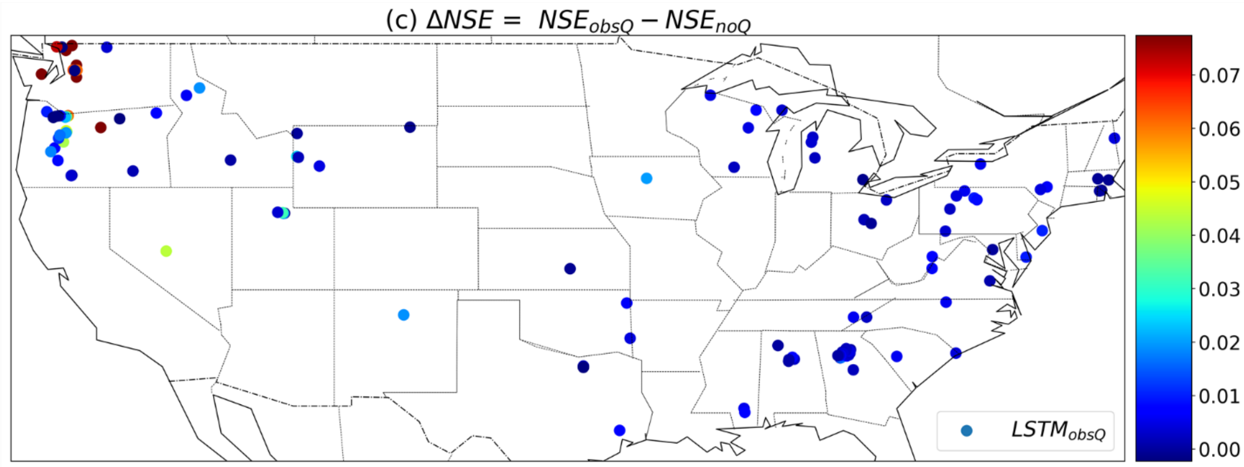


Figure 4. NSE difference between $\text{LSTM}_{\text{obsQ}}$ and LSTM_{noQ}

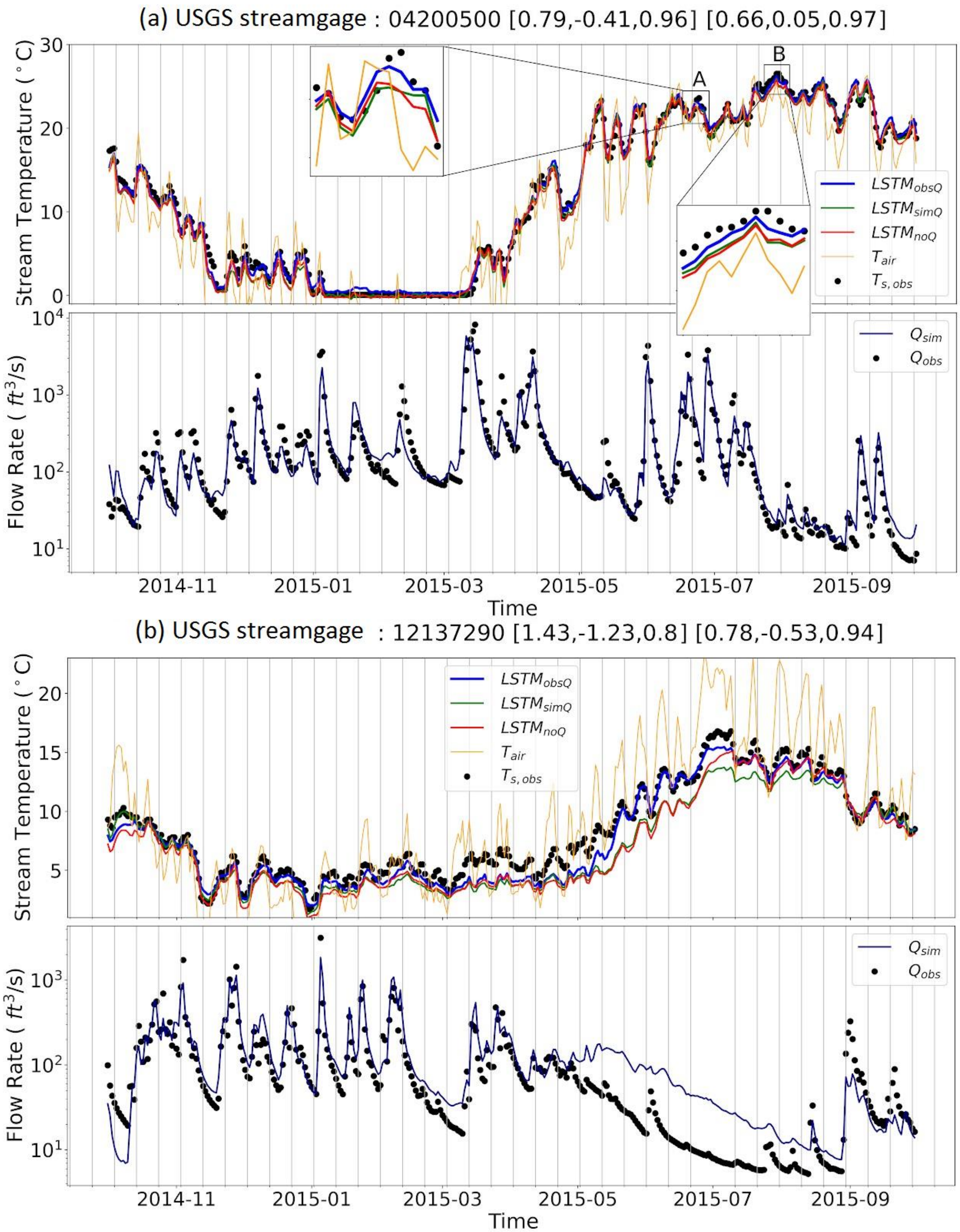


Figure 5. Time series plots of observed and simulated T_s in (a) Black River at Elyria, Ohio; (b) South Fork Sultan River, Washington. The two brackets contain values for [RMSE, Bias, NSE_{res}] for $LSTM_{noQ}$ and $LSTM_{obsQ}$, respectively.

Results

- Both obsQ & simQ are beneficial
- Median RMSE for $LSTM_{obsQ}$: 0.69 ℃. %20 better than $LSTM_{noQ}$
- Median RMSE for $LSTM_{simQ}$: 0.81 ℃, still lower than $LSTM_{noQ}$
- Indirectly absorbing discharge data (through a trained discharge model, i.e. simQ) can be useful if discharge observations (obsQ) are not available.
- obsQ helps more in western part of the US
- Negative biases with $LSTM_{noQ}$ were attributable to undershooting T_s peaks in both winter and summer in some sites (e.g., Figure 5a) and a more consistent bias at other sites (e.g., Figure 5b).
- T_s peaks are often associated with streamflow peaks (possibly caused by warm rain) in the winter but after-storm recession limbs in the summer.

4- DATA AVAILABILITY & RESERVOIRS' IMPACT ON PREDICTION ACCURACY

Based on the previous results, we favor LSTM for stream temperature simulation. However, there are open questions regarding how this model performs when we move into intermittently-measured or unmeasured basins, how performance changes when we encounter reservoirs, and how to assemble the most suitable training data for them.

Data Availability

Here, we created three different data availability groups (DAG, with varying sampling frequency):

a) 415 basins with more than 10% observed T_s data in both training and testing period

b) 306 basins with more than 60% T_s data

c) 99 basins with more than 99% T_s data

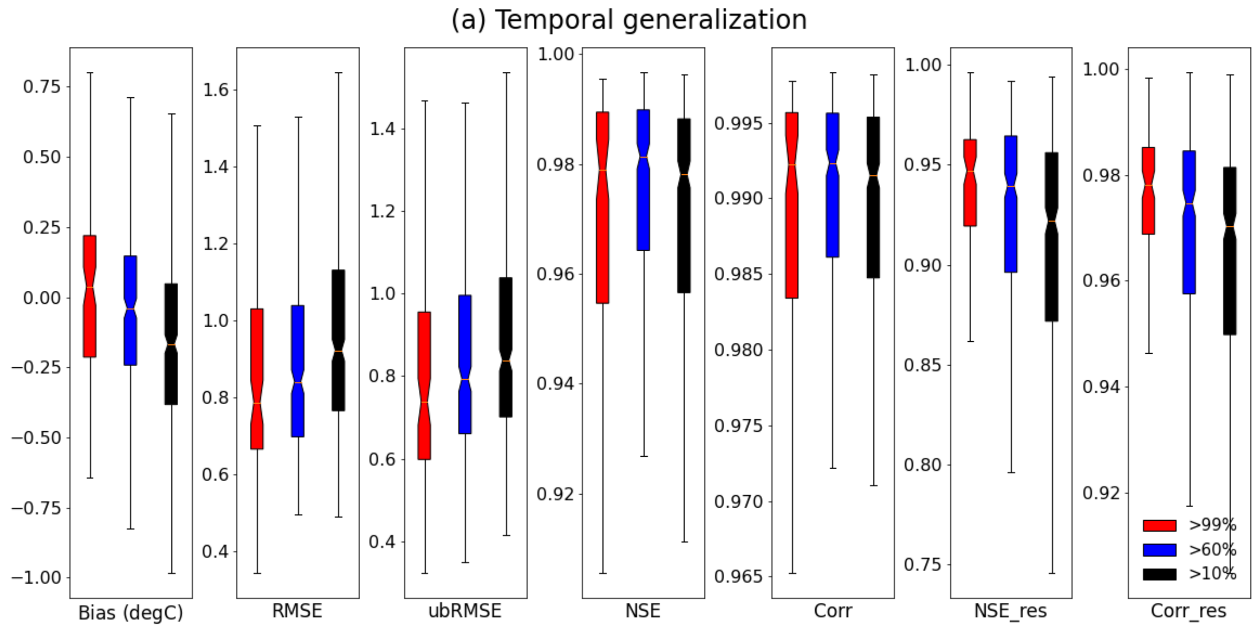


Figure 6. Stream water temperature results for different data availability groups in testing period

Reservoirs Impact

To more explicitly investigate reservoir impact on water temperature, Each of the data availability training datasets was divided into two sub-groups:

(a) basins with at least one major dam upstream

(b) basins without any major dams upstream

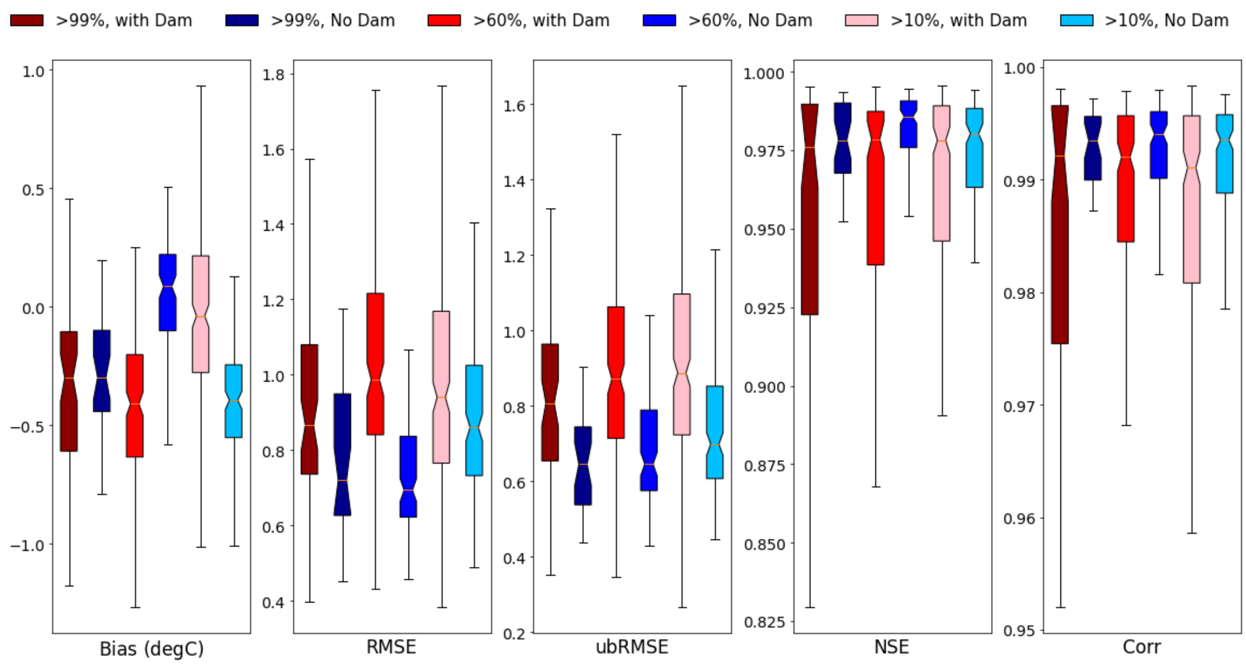


Figure 7. Reservoir impact on stream water temperature model

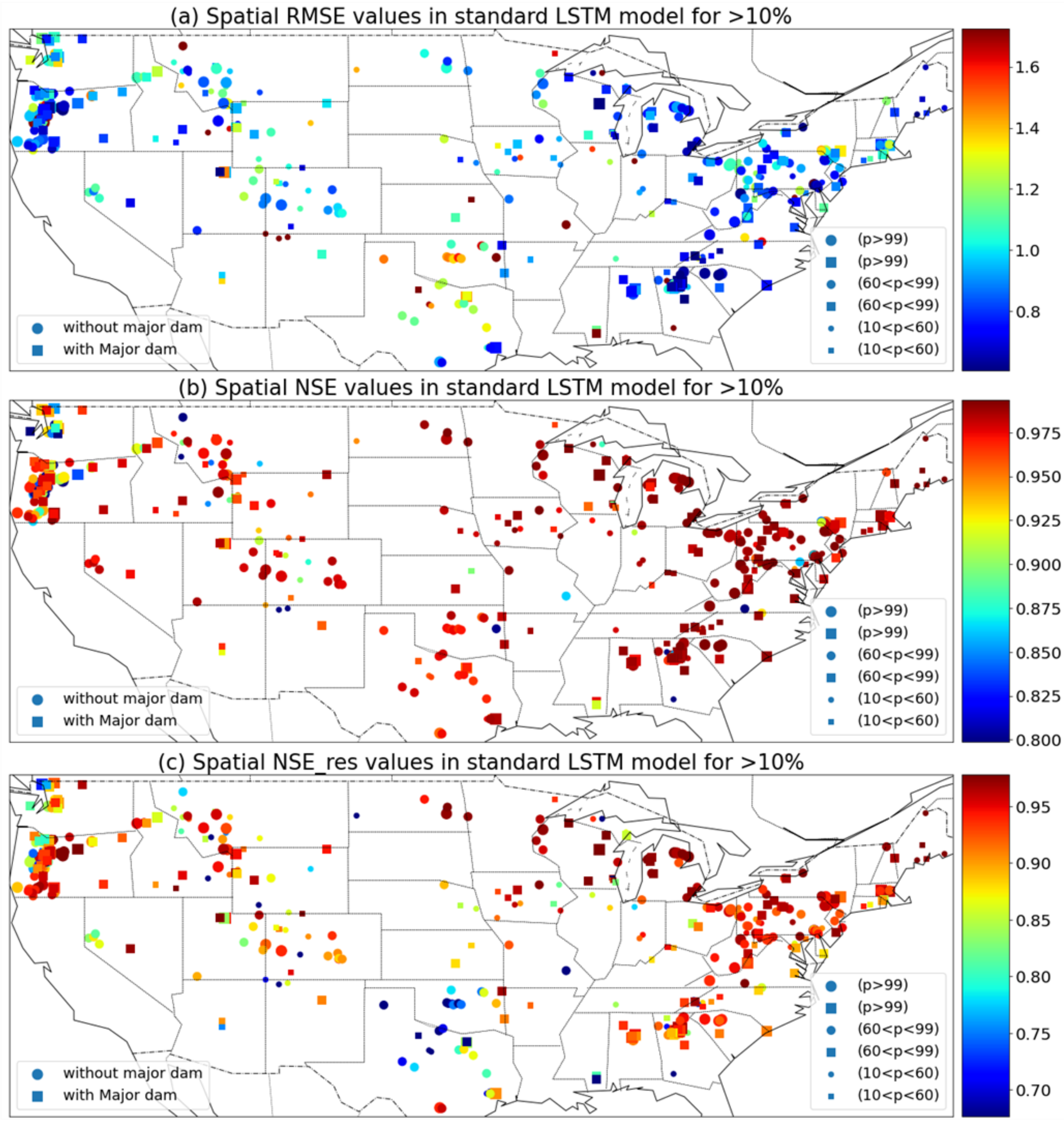


Figure 8. Spatial distribution of RMSE, NSE, & NSE_{res} for %10 DAG

Results

- More intermittent data availability, reduces the accuracy of the results
- Reservoirs adversely affect the accuracy of the T_s model, all of our models understimate water temperature for basins with major dams.
- Major dams increase ubRMSE by an average of >0.2 ℃, and enlarged negative bias by an average of > 0.35 ℃ .

- Even for the reservoir group, the RMSE of the LSTM model appears to be smaller than those reported in other studies.

5- PUB TESTING & INPUT SELECTION ENSEMBLE IMPACT ON RESULTS

Prediction in ungauged basins (PUB)

Models trained on different DAGs were tested with out-of-training-set test monitoring stations, these were those with observations recorded only 10% or less of the time (40 basins).

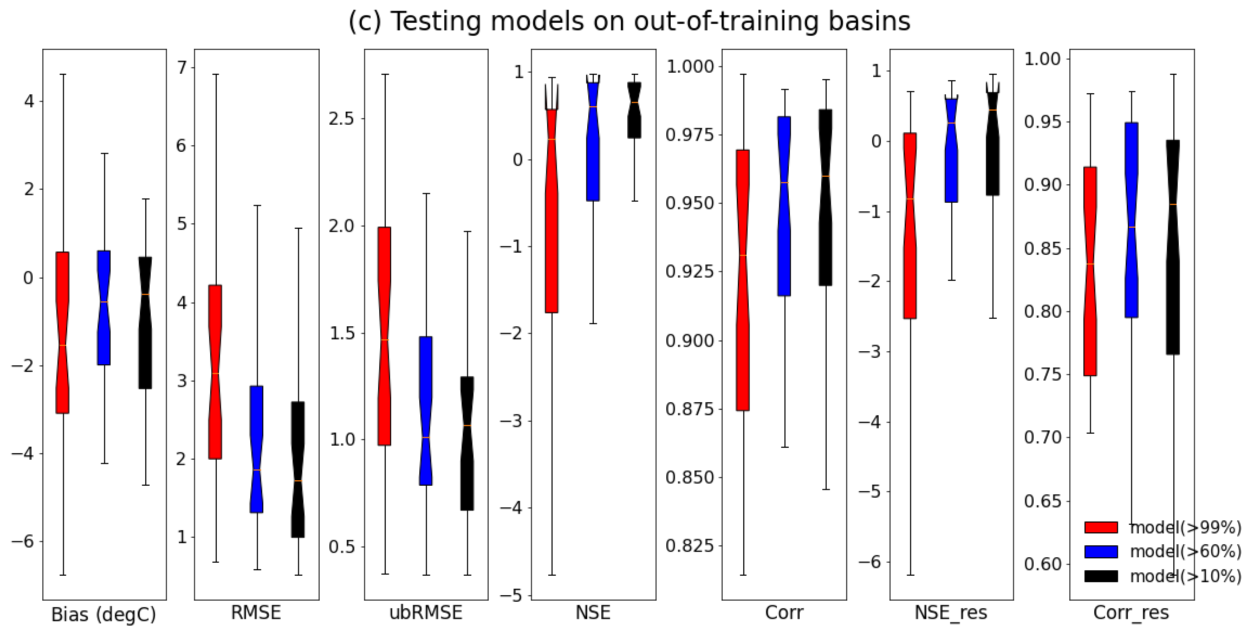


Figure 10. PUB test on 40 basins.

Results

- Correlation and residual correlation still remained high overall, as did the seasonality-removed metrics. Bias was a much bigger issue.
- Better performance when more basins are included (even though they have sparser temperature data) - this trend is opposite that for temporal extrapolation.

- Put more simply, a more diverse training set can cover more regions in the input space, leading to a more robust model.

Input Selection Ensemble

Input selection ensemble means training models with different subsets of the static attributes as inputs, testing each one of the models in ungauged basins and averaging all the results as final prediction.

The full-attribute model setup (full-attr) used all available static attributes, just as in section 4.

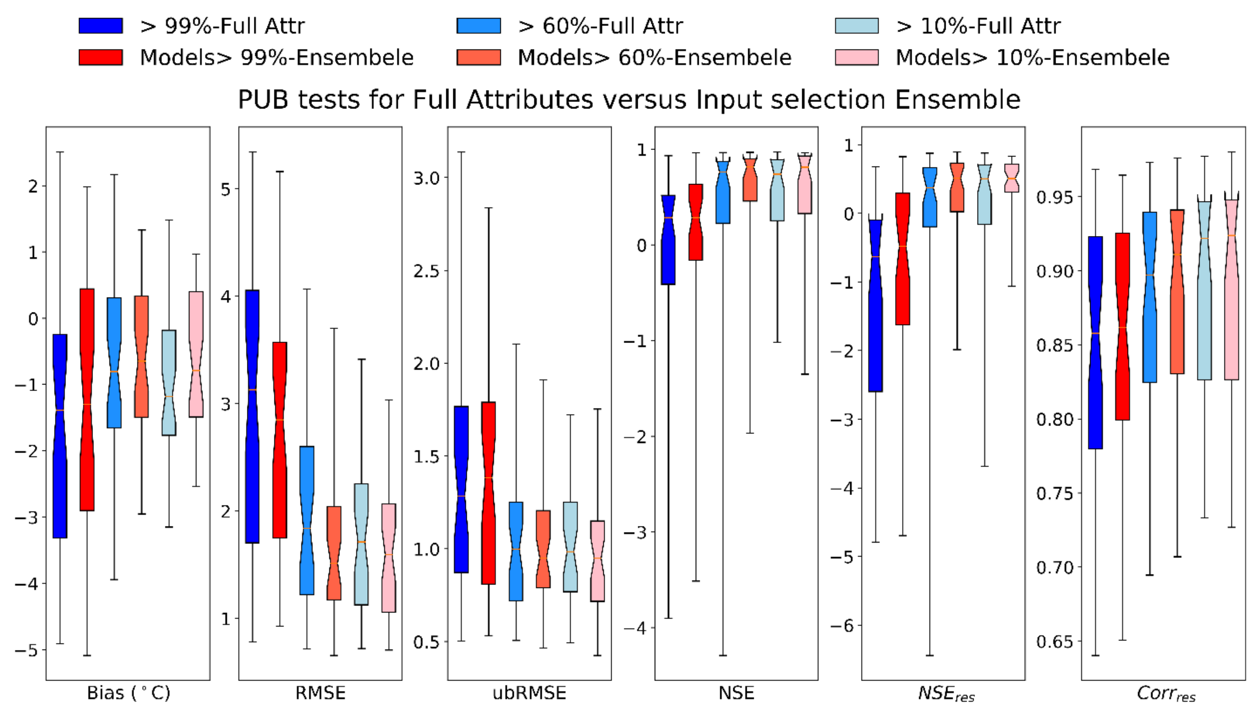


Figure 11. PUB testing and ensemble input selection testing for ungauged basins

Results

- Input selection ensemble (ISE) improved the median of all metrics in all DAGs compared to full-attribute models. We hypothesize that at the ensemble level, ISE reduces overfitting to any one basin attribute, thus improving PUB performance.

Note:

This information is preliminary and is subject to revision. It is being provided to meet the need for timely best science. The information is provided on the condition that neither the U.S. Geological Survey nor the U.S. Government shall be held liable for any damages resulting from the authorized or unauthorized use of the information.

AUTHOR INFORMATION

Farshid Rahmani[1], Kathryn Lawson[1], Wenyu Ouyang[2], Alison Appling[3], Samantha Oliver[4], Chaopeng Shen[1]

ABSTRACT

Stream water temperature (T) is a variable of critical importance and decision-making relevance to aquatic ecosystems, energy production, and human's interaction with the river system. Here, we propose a basin-centric stream water temperature model based on the long short-term memory (LSTM) model trained over hundreds of basins over continental United States, providing a first continental-scale benchmark on this problem. This model was fed by atmospheric forcing data, static catchment attributes and optionally observed or simulated discharge data. The model achieved a high performance, delivering a high median root-mean-squared-error (RMSE) for the groups with extensive, intermediate and scarce temperature measurements, respectively. The median Nash Sutcliffe model efficiency coefficients were above 0.97 for all groups and above 0.91 after air temperature was subtracted, showing the model to capture most of the temporal dynamics. Reservoirs have a substantial impact on the pattern of water temperature and negative influence the model performance. The median RMSE was 0.69 and 0.99 for sites without major dams and with major dams, respectively, in groups with data availability larger than 90%. Additional experiments showed that observed or simulated streamflow data is useful as an input for basins without major dams but may increase prediction bias otherwise. Our results suggest a strong mapping exists between basin-averaged forcings variables and attributes and water temperature, but local measurements can strongly improve the model. This work provides the first benchmark and significant insights for future effort. However, challenges remain for

basins with large dams which can be targeted in the future when more information of withdrawal timing and water ponding time were accessible.

REFERENCES

Delpla, I., A.-V. Jung, E. Baures, M. Clement, and O. Thomas. 2009. “Impacts of Climate Change on Surface Water Quality in Relation to Drinking Water Production.” Environment International 35 (8): 1225–33. https://doi.org/10.1016/j.envint.2009.07.001.

Justice, Casey, Seth M. White, Dale A. McCullough, David S. Graves, and Monica R. Blanchard. 2017. “Can Stream and Riparian Restoration Offset Climate Change Impacts to Salmon Populations?” Journal of Environmental Management 188 (March): 212–27. https://doi.org/10.1016/j.jenvman.2016.12.005.

Kaushal, Sujay S, Gene E Likens, Norbert A Jaworski, Michael L Pace, Ashley M Sides, David Seekell, Kenneth T Belt, David H Secor, and Rebecca L Wingate. 2010. “Rising Stream and River Temperatures in the United States.” Frontiers in Ecology and the Environment 8 (9): 461–466. https://doi.org/10/bfxq26.

Madden, N., A. Lewis, and M. Davis. 2013. “Thermal Effluent from the Power Sector: An Analysis of Once-through Cooling System Impacts on Surface Water Temperature.” Environmental Research Letters 8 (3): 035006. https://doi.org/10.1088/1748-9326/8/3/035006.

Marcogliese, David J. 2001. “Implications of Climate Change for Parasitism of Animals in the Aquatic Environment.” Canadian Journal of Zoology 79 (8): 1331–52. https://doi.org/10.1139/z01-067.

Martins, Eduardo G., Scott G. Hinch, David A. Patterson, Merran J. Hague, Steven J. Cooke, Kristina M. Miller, David Robichaud, Karl K. English, and Anthony P. Farrell. 2012. “High River Temperature Reduces Survival of Sockeye Salmon (Oncorhynchus Nerka) Approaching Spawning Grounds and Exacerbates Female Mortality.” https://doi.org/10.1139/f2011-154.