Climate Change Impacts on the Occurrence of Riverine Heatwaves

Tristan Ballard
Department of Earth System Science
tballard@stanford.edu

***Abstract*—River warming in response to climate change may have far-reaching ecological and socioeconomic consequences. Unfortunately, large temporal gaps in the instrumental record limit the ability to study these systems. To address this, I model daily river temperatures across the U.S. using a long short-term memory neural network with the goal of gap-filling historical records, and I then estimate historical shifts in the probability of extreme heat events. The model achieves high accuracy, with a median R^2 of 0.91 at locations not used to train the model. Results suggest that riverine heatwaves have already increased in duration across much of the U.S., with trends likely to accelerate in the future.**

***Keywords*—river temperature, LSTM, recurrent neural networks, climate change**

I. Introduction

Water temperature is a fundamental property of aquatic ecosystems.[1] Riverine water temperatures directly impact the physiological processes of fish and other organisms, determining their metabolic rates, growth, life cycles, and habitat suitability. Temperature can also indirectly influence water quality, with warmer waters increasing cyanobacteria growth, denitrification rates, and contaminant toxicity.

Moreover, climate extremes play a critical role in structuring terrestrial and marine ecosystems.[1,2] One such climate extreme, atmospheric heatwaves, has become considerably more frequent and intense in recent decades in response to climate change, and recent work has shown that marine environments, too, are already experiencing prolonged and more intense heatwaves. Evidence, or lack thereof, that climate change has already impacted freshwater temperatures is relevant to a variety of applications, including ongoing or future litigation regarding climate change and thermal pollution. For instance, fisheries and conservation groups recently sued the Environmental Protection Agency to protect salmon in the Pacific Northwest, arguing in part that the salmon are threatened by warming rivers. Further, if climate change has already impacted observed riverine heatwave (RHW) trends, then it is reasonable to assume that trends will continue into the future, albeit at a potentially accelerated rate.

While there are minimal field or experimental studies on the ecological effects of riverine heatwaves, there is compelling evidence that aquatic ecosystems are vulnerable to prolonged, extreme heat events.[2] Marine heatwaves, for example, have been linked to both reduced and increased algal growth, re-distribution of fish species, disease outbreaks, widespread coral bleaching, and mass mortalities. So to identify riverine ecosystems most vulnerable to climate risks, it is crucial to analyze and track RHWs both historically and into the future.

However, no study to date has analyzed changes in RHWs, due, in part, to a lack of consistent daily-scale temperature data. Large temporal gaps in the river temperature instrumental record (from, for example, sensor malfunction) limit the ability to calculate long-term heatwave metrics. Here I address, for the first time, how RHWs have changed in recent decades by developing a model capable of gap-filling existing records.

I address the riverine temperature gap-filling challenge by testing several neural network architectures on a 37-year record of daily measurements taken on the Delaware River, subsequently applying the final model to estimate river temperatures and relevant heatwave metrics at 253 U.S. rivers. I find that a long short-term memory recurrent neural network performs best on the selected training data. Further work will investigate expanding the model to train on more riverine data, incorporate additional input features, and predict riverine temperature and RHWs responses to future climate change.

II. Data

A. Response: Riverine temperature

The U.S. Geological Survey (USGS) maintains a network of thermometers in rivers across the U.S., providing daily temperature measurements.[3] To coincide with input feature availability, I therefore restrict the analysis to the years 1981 to 2017. I also select only those river sampling locations with at least 200 temperature measurements and only rivers determined by the USGS to have had minimal human influence, such as from artificial dams, and are therefore suitable for studying the impacts of climate variables. This results in 254 sites (Figure 1). I ultimately train and evaluate the model on only one of these sites, the Delaware River, since it has minimal missing values.

B. Input features: Air Temperature and Precipitation

Air temperature and discharge have been identified as two key drivers of riverine temperature. Shallow, low-flow rivers, for example, are highly sensitive to fluctuations in air temperature. As a proxy for discharge, for which there is not widely available data, I use precipitation. For both air temperature and precipitation, I use 4km resolution daily data from the PRISM climate group[4], aggregated over each river's watershed boundary (Figure 1).

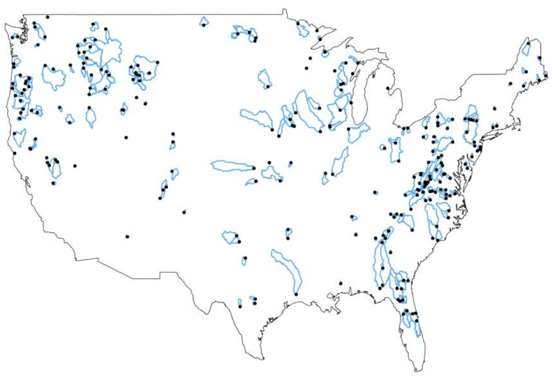

Figure 1. River temperature measurements are available for 254 rivers (points; watershed boundaries in blue) across the U.S. I train my model on one of these, the Delaware River, due to its minimal missing data.

III. Candidate models

The Delaware River data used to train and evaluate candidate models exhibits clear seasonality and autocorrelation (Figure 2). This suggests that recurrent neural networks may prove particularly effective, since they can flexibly incorporate information from previous states to inform the current value's prediction. I use mean-squared error (MSE) as the target function.

I separate the 37-year data record (n=12,628) into training (70%), evaluation (20%) and test (10%) sets. I partition these chronologically rather than randomly in part out of convenience. If there were clear evidence of a long-term, decadal trend, or particularly anomalous years (the test set *does* seem to be cooler on average), I would be less confident in this partitioning and would opt for a more randomized implementation.

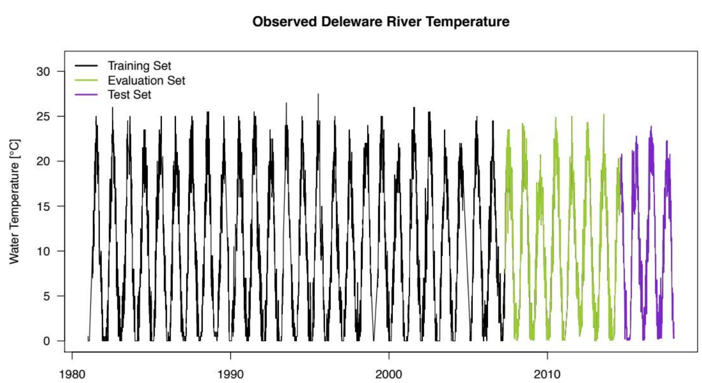

Figure 2. The Delaware River data (n=12,628) used to build the model partitioned into training (70%), evaluation (20%), and test (10%) sets.

A. Baseline Error

It is helpful to have a baseline error to compare model performances against. I therefore consider a 'naïve' MSE to be the MSE resulting from predicting today's water temperature as the average of the air temperature for the previous 7 days. Statistical models should, if the input features are indeed relevant, be able to outperform this naïve metric.

B. Neural Networks

I consider artificial neural networks, and two types of recurrent neural network formulations: gated recurrent units (GRUs) and long short-term memory (LSTM) models. GRUs and LSTMs are similar, with GRUs generally sacrificing some accuracy relative to LSTMs for lower computational cost, a common tradeoff in machine learning applications.

For all, I consider different combinations of dropout regularization (for recurrent networks there are two forms of dropout one may include for each layer), Adam vs. RMSprop optimizers, and different numbers of layers, input dimensionality, training epochs, and minibatch sizes. I dedicated comparatively low time to tuning parameters for the ANN, so its lower performance may be in part due to personal bias.

For the GRU and LSTM models, I use a lookback period of 14 days, meaning that the model can *directly* take into account only data from the previous 14 timesteps in estimating the current value. I did not tune this parameter.

I standardize all three variables by centering and dividing by the respective standard deviations prior to fitting the models. MSEs in Table 1 are in standardized units.

IV. Results

The LSTM model performed best on the evaluation set. It has dropout=0.1 and recurrent dropout=0.1. Because there are only two additional input features and a lookback period of only 14 timesteps, along with a large training set, I think it was difficult for the LSTM model to overfit the data, resulting in

Model	**Architecture**	**Evaluation Set (n=2,600) MSE**
Naïve baseline	Average 7-day air temperature	0.133
ANN	Flattened input layer (outputDim=84), dense layer (outputDim=32), dense layer (outputDim=1)	0.069
GRU	GRU layer (outputDim=16), dense layer (outputDim=1)	0.059
CNN + GRU	1D convolution layer, maxPooling, 1D convolution layer, GRU layer, dense layer	0.059
LSTM	1 LSTM layer (outputDim=16) followed by two dense layers (outputDim=1)	**0.055**

improved performance at lower levels of dropout regularization. I also ran the model for 15 epochs and batch size of 32. The model performs well on the held out test data (n=1,228) (Figure 3) with an R^2 of 0.95. The Adam optimizer tended to perform better than the RMSprop algorithm. There is also an interesting dip in observed temperature in the summer of 2015 that the model failed to capture (Figure 3), suggesting perhaps that something aside from air temperature and precipitation may have caused the decrease.

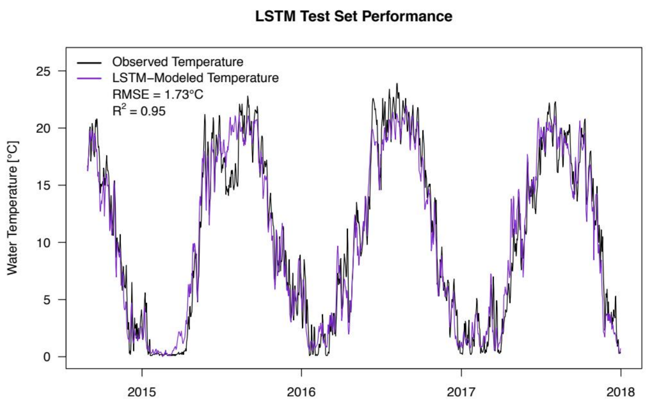

Figure 3. The final LSTM model achieves high accuracy on the test set, capturing both the magnitude and seasonality of temperatures.

V. MODEL APPLICATION

Given that the LSTM model perfoms well on the test set for the Deleware River site, I was interested to see how the same model perfomed when applied to the remaining 253 sites identified previously (feature inputs appropriately standardized prior to applying the model). Despite these rivers coming from distinct regions—where the relationships between air temperature, precipitaiton, and water temperature are not expected to be identical—the model performs well, with a median R^2 of 0.91 and median RMSE of 2.1°C. The RMSE would likely improve had training data from these sites been incorporated as well into the model build.

With this reassuring performance on the remaining river sites, I then take the fitted daily riverine temperatures at each location and estimate the occurrence of RHWs. I use a simple metric for RHWs used occasionally in atmospheric heatwave studies, namely exceedance of the historical 95th percentile temperature for a location, for at least three consecutive days. I then calculate the linear trend over the historical 1981-2017 period at each site (Figure 4), which indicate that there have been predominantly increasing trends in the number of RHW days, at an increase of 20-30 days per decade at several sites. Interestingly, the upper Midwest indicates decreasing likelihoods of RHWs.

VI. FUTURE DIRECTIONS

I would like to incorporate training data from all stations rather than a single station. The Deleware River, with 12,628 observations, is an excellent case study, but the combined measurements across the 254 sites is roughly 500,000, suggesting a great opportunity for expanding the generalizability of the model, and conceivably supporting a more complex model. I was also nearly able to incorporate simulations of air temperature and precipitation from a climate model to predict river temperature in the coming decades across the U.S., and, potentially, the globe, a key end-goal of this analysis. Further, online tutorials suggested differencing the data, due to its high autocorrelation, may improve performance. Attention-based model structures may also better incorporate measurements from thousands of previous timesteps (e.g. temperature on the same date for multiple previous years) than the GRU/LSTM formulations.

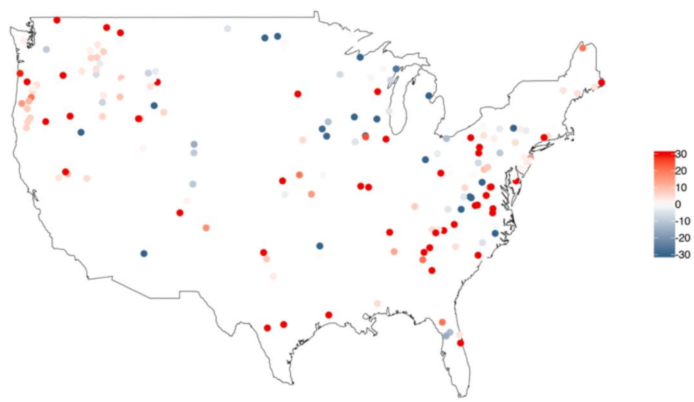

Figure 4. I apply the final LSTM model to estimate daily river temperatures at the remaining 253 sampling locations. The model performs well on these locations, with a median RMSE of 2.1°C and R^2 of 0.91. Then, for each sampling location, I estimate the decadal trend in heatwave days (points), defined as days where river temperatures exceed the local historical 95percentile for at least 3 consecutive days. The majority of stations have seen increases in the number of heatwave days. Units are change in number of heatwave days per decade between 1981 and 2017.

ACKNOWLEDGMENT

I would like to thank Sarah Najmark for her guidance throughout on the project. Code is available at https://github.com/tristanballard/RiverTemperatureNeuralNetworks.

REFERENCES

[1] Comte and Olden. Climatic vulnerability of the world's freshwater and marine fishes. Nature Climate Change (2017).

[2] Frolicher *et al*. Marine heatwaves under global warming. *Nature* (2018).

[3] USGS Water Data for the Nation. waterdata.usgs.gov/nwis/qw

[4] PRISM Climate Group, Oregon State University, prism.oregonstate.edu.