Short-Term River Forecasting with a Stacked Ensemble of Tributary Models

Patrick J. Donnelly
Department of Computer Science
Oregon State University-Cascades
Bend OR, United States
patrick.donnelly@oregonstate.edu

Orion Junkins
Department of Computer Science
Oregon State University-Cascades
Bend OR, United States
junkinso@oregonstate.edu

2022 7th International Conference on Frontiers of Signal Processing (ICFSP) | 978-1-6654-8158-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICFSP55781.2022.9924877

***Abstract*—The ability to accurately forecast river levels is a difficult task. Fluctuations in the flow of each river are determined by complex and localized weather, topological, and geographic factors, and it is challenging to generalize a single hydrologic model across different rivers. Furthermore, researchers struggle to best extract the trends among a large amount of measurement data. We propose an ensemble of recurrent neural network models to forecast river flow using weather data sampled at tributaries along the catchment. We model weather sites for 12 tributaries along the catchment to capture each site's contribution to the flow. These individual models are used to train a second-stage ensemble model to predict overall river flow. We explore our forecasts and results on three rivers in the Pacific Northwest of the United States.**

***Keywords*—deep learning, ensemble, hydrology, river flow**

I. Introduction

Fluctuations in river flow rate are impactful to the many different stakeholders that live, work, and play nearby. These include landowners who live adjacent, agricultural interests that rely on the water, and those who recreate nearby. Furthermore, accurate river level prediction is necessary for water resource management, drought mitigation, and flood prediction. Rapid fluctuations in flow challenges efficient water management and creates dangerous situations for both life and property.

Although changes in river flow are closely correlated to a combination of short and long-term weather patterns, there are many other complex contributing factors. A river's rate of flow is influenced by the size of the watershed, topography, rate of snowpack melt, number and variety of contributing streams, soil composition, geology and the amount and type of vegetation cover. Together these interrelated factors create a complex and non-linear hydrological system which has proven difficult to model with traditional statistical methods.

Neural networks are well-suited for such a task and numerous prior studies have explored this potential [1]. However, the amount of data needed for accurate forecasting with deep-learning is often restrictive. The high degree of interrelated data challenges the accuracy, generalizability, and real-world utility of computational models predicting river flow.

Prior studies have approached this curse of dimensionality in two different ways. In one general approach, researchers attempt dimensionality reduction techniques, such as principle component analysis, clustering, feature selection, selective sampling, or filtering from the input space [2]. These approaches have the benefit of smaller network topologies, shorter training times, and fast inference times. However, reductions in the feature space often hinder accuracy of such models. In the second approach, researchers attempt to use all available data. These approaches often result in large model architectures, long training times, and slow inference times. For these reasons, these models are often challenging to deploy at scale or in real time.

A. Related Work

Traditional approaches to river flow prediction include numerical, physical, and data-driven models, and researchers have attempted to ensemble different statistical approaches together [3]. Neural networks have been used to model hundreds of different river systems (see [2] for review), although these studies vary greatly in their choices of model architecture and input feature space. Recently, approaches using deep-learning have significantly outperformed older approaches using other traditional machine learning techniques (see [1] for review).

Convolutional neural network (CNN) [4] and Long Short-Term Memory (LSTM) [5] architectures are most often tasked for this problem. Recent approaches have explored other approaches, such hybridizing CNN and LSTMs together [6] or adapting models from other fields to hydrology [7]. In all approaches, researchers in hydrologic forecasting are challenged to make the best use of a large amount of measured data [8].

B. Contributions and Novelty

This paper presents a stacked ensemble of deep learning models to predict river flow. Specifically, we train individual models at the level of the individual tributary and use these intermediate predications to train a second-stage meta-model. This approach enables each individual base model to attempt to model the tributary's contributing effect on the overall flow. The ensemble combines the predictions of the individual models to produce an overall prediction of the river flow. We demonstrate the accuracy of our approach through a case study of the Illinois River in southwestern Oregon and we compare our results with NOAA's Advanced Hydrologic Prediction Service tool.

Our approach has several important advantages. First, we leverage all available data across the catchment to maintain accuracy without compromising the potential for practical deployability. In effect, we achieve a dimensionality reduction by first individually modeling smaller areas of the overall catchment before combining these models to predict the overall river flow. Second, our approach results in smaller individual models that are fast to train, easy to parallelize, and quick to perform inference. Compared to many other deep learning approaches, our approach has the advantage of enabling the system to be deployed in delayed real-time.

Third, our approach facilitates modularity in choice of deep learning models included in the pipeline. This enables the individual tributaries to be modeled by any time-series models and different model architectures can be used to model different tributaries. Lastly, our ensemble approach provides an added level of transparency that could aid in interpreting the overall results of the model. While the constitute deep learning models suffer from the lack of interpretability associated with deep learning, our approach provides predictions of each tributary-level model which allows examination of which tributaries may be disproportionately contributing to a particular fluctuation.

This paper describes the data chosen to train our models, the processing methods applied to that data, and the model architecture used to produce forecasts. A case study provides measured accuracy of the final model along with a quantified comparison against the best currently available alternative forecast provided by the National Weather Service[1]. To provide a more intuitive presentation of these results, forecasts are visualized over time. We conclude with a discussion of results, limitations, and the potential for real world deployment.

II. Dataset

No two river systems are the same, varying greatly in their catchment size and topology. Geological factors, such as soil composition, greatly affect the amounts of ground saturation and runoff. Rivers also vary in the types of in-flow sources, such as springs, rain, glacial melt, or snowmelt. River flow rate can also be greatly affected by dams or other human intervention.

Modeling river flow rate requires access to flow data for the river itself and weather data for many relevant locations throughout the drainage. In this work, we focused on free-flowing rivers that are only affected by natural water sources. We restricted to free-flowing rivers because dams are controlled by humans and their flow is not the direct result of weather.

A. *Site Locations*

We selected three rivers in the Pacific Northwest region of the United States, chosen primarily for the availability of comprehensive historic weather data. These rivers have small to medium sized catchments, do not contain a dam, and feature a diversity of weather patterns, including rain, snow and extended dry periods. All three rivers are popular recreational sites for both public and commercial use. The three chosen rivers vary in size, elevations, and water sources. We hope that demonstrating efficacy with this set of rivers will imply our ability to scale to any other catchments with similar conditions.

Icicle Creek is tributary of the Wenatchee in central Washington. It is around 32 miles (51 km) long and has a catchment of 212 square miles (551 km^2), fed by rain and snow. With elevations over 5000 feet (1500 m), this is a high elevation catchment and has snowmelt most of the year. We modeled the USGS gauge known as *Above Snow Creek*.

White Salmon River is a tributary of the Columbia River in northern Oregon. It runs for 44 miles (71 km) with a catchment of around 400 square miles (1036 km^2). This catchment is at moderate elevations and is fed by snow melt, rain, and natural springs. It receives year round snow and glacier melt. We modeled the USGS gauge known as *Near Underwood*.

The Illinois River (Oregon) is a tributary of the Rogue River in southwestern Oregon. It is a moderately sized catchment of 983 square miles (2550 km^2). This catchment is further south than the other two rivers and at a lower elevation. Although fed by rain and snowpack, snow melt only occurs for a small part of the year. We modeled the USGS gauge known as *Near Kirby*.

B. *Flow Rates from USGS*

The United States National Geological Survey (USGS) provides river flow data measured in ft^3/s for numerous river sites across the country. These gauges provide hourly flow rates for many sites across the country. All of the sites were chosen because they have a USGS gauge with an extensive record of historical data available (see TABLE I.). These flows serve as the ground truth we attempt to model in our experiments.

TABLE I. USGS Gauges Modeled

Gauge Name	*Gauge ID*	*Data Since*
Icicle above Snow Creek	12458000	1993-10-01
White Salmon near Underwood	14123500	1987-10-01
Illinois near Kerby	14377100	1988-10-01

C. *Features from OpenWeatherMap*

We collected weather data using the OpenWeatherMap API[2]. This tool provides both historical weather data and current forecasted weather data for any latitude and longitude. We chose 12 sites evenly distributed throughout the tributaries of each river. Of the weather values provided, we retained temperature, rainfall, snowfall, and humidity as features. We then generated seven additional features, including the day of the year to serve as a proxy for season. We also calculate snow accumulation, rain accumulation, and average temperate for 10 and 30 day windows. This gives use a total of 11 features per weather location. We normalized all features in range $[0..1]$ and imputed missing values with the daily average.

III. Methiods

We selected 11 features per weather site. We model 12 weather sites per catchment, resulting in an input space of 132 features per time sample. We sample data hourly across five days resulting in 120 samples, yielding 15,840 values across the sequence. Large input spaces necessitate extensive computational resources to train and high order relationships between the data can be lost amidst the complexity. We propose to mitigate this problem by training a single tributary model on each individual weather site. We combine individual model predictions through an ensemble to predict the overall flow rate.

[1] https://water.weather.gov/ahps/forecasts.php
[2] https://openweathermap.org/api

A. Tributary Forecasting Model

As a first stage, we model how a single specific site affects the river flow rate. This model takes the recent river level, recent weather history, and the forecasted weather for a total of 11 features. These models are then trained to predict the river flow at a gauge downstream. Fig. 1 illustrates this pipeline for a single tributary forecasting model. This is a flexible and modular pipeline. It supports any time-series model, and different tributaries cans be modeled with different architectures.

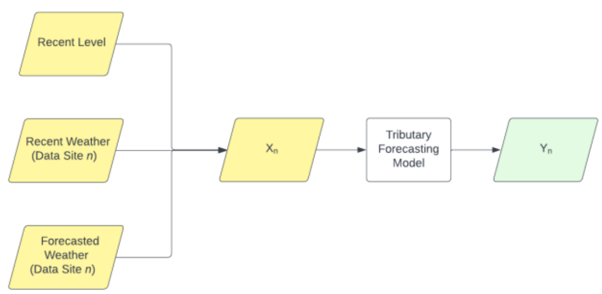

Fig. 1. Diagram of an individual tributary forecasting model, where X_0 represents the input features and Y_0 represents the tributary-level prediction for forecast $site_0$.

The individual tributary forecast model serves to capture the complex high order relationships between the weather features localized at a specific geographic site. Each of these models are specific to one single weather monitoring site and we must train an individual model for each of the twelve sites. However, each individual tributary model is lightweight and efficient to train.

While these individual tributary forecasting models often achieve reasonable levels of accuracy, they also suffer from high degrees of variance in their efficacy. This is expected as localized weather systems impact some sites more than others. Any individual site only makes small contributions to the river flow and cannot predict contributions of other unknown sources.

Fig. 2. Diagram of the ensemble model that combines n individual tributary forecasting models to produce a single prediction of the river flow.

B. Ensemble Model

Next, we leveraged the predictions of the individual tributary forecast models in a second stage model. This ensemble approach, shown in Fig. 2, takes the 12 intermediate predictions and learns a single prediction of the river level. By ensembling tributaries, we reduce the dimensionality of the feature space while retaining contributions from all original data.

IV. Results

We trained a Block Recurrent Neural Network forecasting model[3] that supports past covariates for a fixed input chuck before prediction time. We compared a basic recurrent neural network, long short term memory network, and a gated recurrent unit (GRU). We selected the GRU for use in our forecasts based on the results of preliminary experiments.

A. Case Study: Illinois at Kerby

We trained our ensemble to predict the historical forecast for the Illinois River (Oregon) at Kerby. Predicting a 24-hour advance forecast window, our system forecasted with an 8.57% mean absolute percentage error (MAPE) over the entirety of the available data. This was a reduction of 10% MAPE from the best performing individual single tributary by itself.

To compare our system to other forecasting models, we require a history of past forecasts. Although such a forecast is available in real-time for a small number of rivers, the past forecasts are not archived. We collected daily forecasts over a period of 110 days from the National Weather Service[4]. Over this sample, our 24-hour forecast model achieved a MAPE of 13.62% compared to 16.47% by the NOAA model.

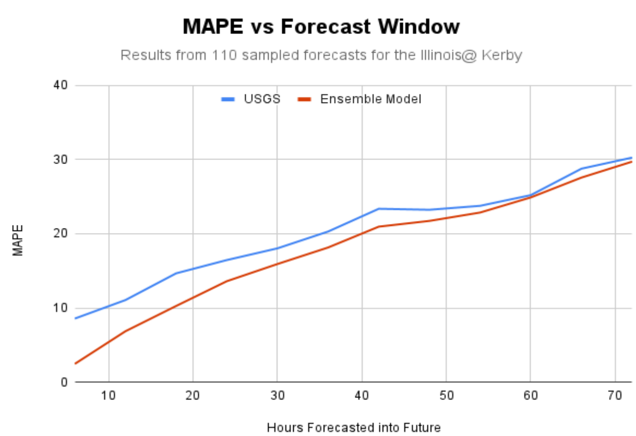

Fig. 3. Comparison of our approach (red) to the NOAA forecast (blue) over different forecasting windows for a period of 110 days in Spring 2022.

Next, we investigated how our model compared to the NOAA forecast as the size of the forecasting window varied, as shown in Fig. 3. We found that our model is much more reactive than their model and thus able provide much better predictions for short forecast windows representing the near future. As the size of the forecasting window increased, the accuracy of our model and NOAA's predictions were similar.

Although this comparison represents an anecdotal analysis of a single river forecast over a single three-month observation, we find these results encouraging and demonstrate the potential of our approach to provide timely and accurate river forecasts.

[3] https://unit8co.github.io/darts/generated_api/darts.models.forecasting.block_rnn_model.html

[4] https://water.weather.gov/ahps/forecasts.php

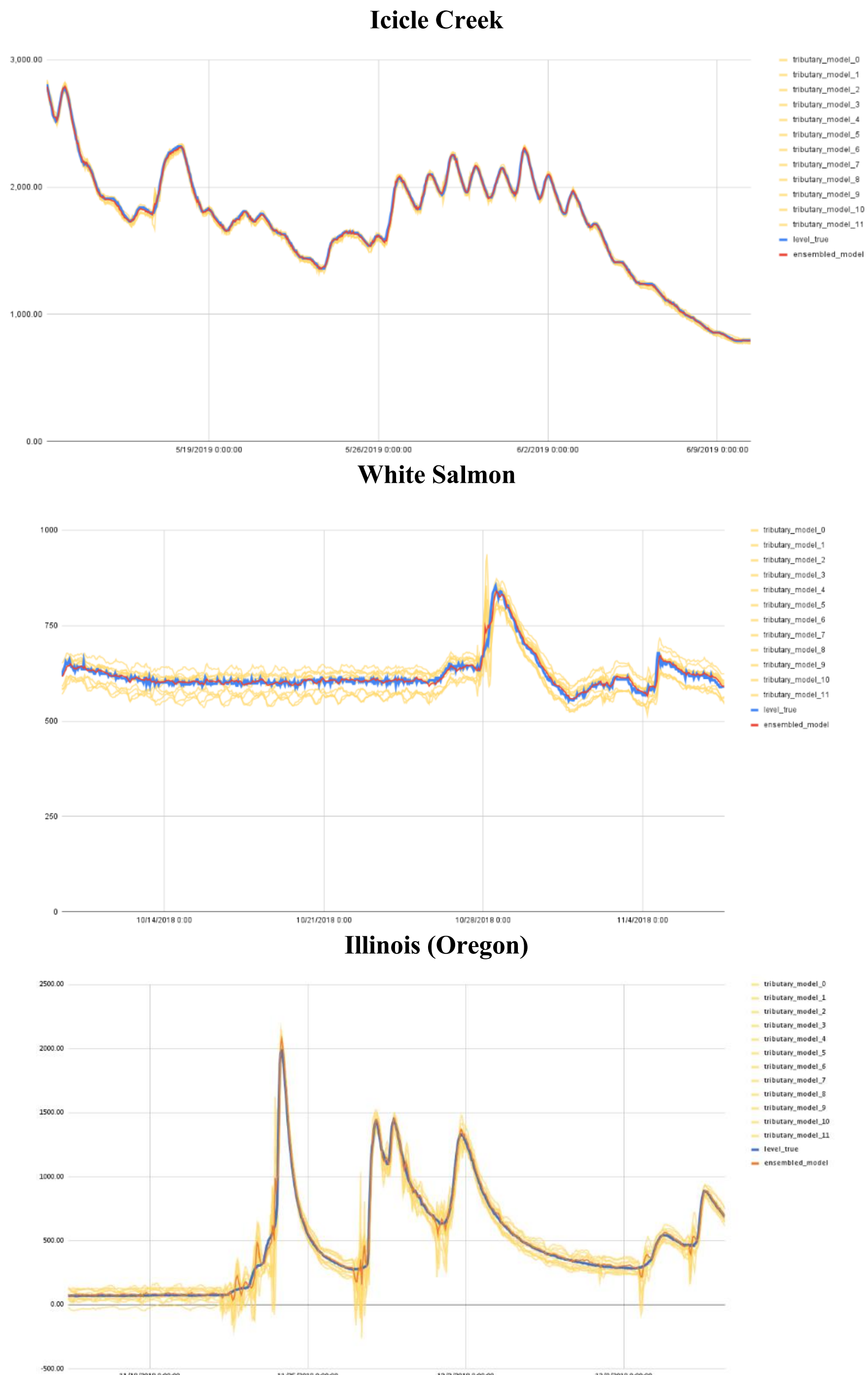

Fig. 4. Three examples of one month of 24-hour forecasts made for Icicle Creek (top), White Salmon River (middle), and the Illinois River (bottom). Twelve tributary level predictions (yellow) are ensembled to predict river flow (red) and compared to the true flow rate (blue).

B. *Visualization of Forecast*

To better illustrate the smoothing effects of our ensemble approach compared to each contributing tributary forecast models alone, we visualize example forecasts as Fig. 4. For each river, we show one month of our tributary level predictions compared to the ensemble prediction and the true flow rate, again using a 24-hour forecasting window.

In the Icicle Creek example, our individual tributary predictions model the overall flow fairly well at times, but small variances consistently occur for all tributary models. However, together, we note the ensemble is highly successful at forecasting this river during this time period.

Next, we demonstrate a forecast for White Salmon River. Here we observe that our individual tributaries were highly variable and inaccurate in their predictions. Interestingly, we observe many cases in which they forecast a similar trend as the river, but the models were off by a fixed margin of error. This illustrates how the weather patterns that influence one site influence the entire river, but one site alone does not have the information to model the entire system. Despite the variances of individual tributary models, our ensemble is moderately successful at forecasting the river trend over time.

The third example shows the Illinois River (Oregon). Again we see examples of significant error made by the individual tributary models but only minimal error made by the smoothed ensemble model. These moments likely represent rainstorms, which affected different sites differently. These storms produce a subsequent sharp rise in the rate of the entire river. Our individual tributary models are reactive to local conditions at the twelve sites, which results in individual mispredictions, but taken together, the ensemble is able to model the overall trend.

V. Conclusion

We present an approach for modeling river flow using an ensemble of individual tributary models. We demonstrated this approach on three small river catchments in the Pacific Northwest. For all three examples, our approach is able to successfully model the flow, while leveraging all available data.

A. Summary of Contributions

We demonstrate that while an individual weather site cannot model the overall flow of a river, an ensemble can leverage multiple sites in the catchment for a more accurate river flow prediction. Our approach facilities an informed data dimensionality reduction that learns complex correlations and interdependencies through a two-stage ensemble approach.

Our approach has the advantage that individual models representing a single weather site are light-weight and easy to train, compared to other complex models that leverage all data at once. These individual modes can be trained in parallel, distributed over a cluster of computer nodes. This makes deployment of our system realistic and practical. Multiple models can be trained concurrently over hours, as compared to days or months for models requiring high dimensional input.

B. Limitations and Future Work

Although the results presented here validate the feasibility of our approach of ensembling individual tributary models for river flow forecasting, they reflect examples of only three small to medium-sized rivers. More work is needed to validate this approach across a larger set of rivers and to generalize how this approach scales to larger rivers with larger catchments. As future work, we will acquire historical data to perform a systematic evaluation of our approach over a large variety of locations.

Our live system models future river forecasts using weather predictions of the future. However, when we run historical simulations, we lack those weather forecasts made at that past date. Given a lack of freely available archival weather forecast data, we must leverage past weather data as a stand in for future weather forecast when simulating historical predictions. Undoubtedly, weather forecasts are not perfect and any perfect knowledge of future weather biases the learning models. In future work we will compare our models built with forecasted weather data against those with perfect weather history. We seek an approach to simulate historical weather predictions by adding sufficient noise to the actual weather records, such that we can leverage them to run realistic historical forecasts over archival data. We are currently archiving months of daily weather forecasts to enable such an approach.

C. Potential for Near Realtime Deployment

Different river stakeholders value different windows of river forecasts. For example, those drawing water for agriculture tend to value long term forecasts while those seeking to recreate on the river may only be concerned about the flow on a particular day. Recreation river users are particularly concerned about sudden and dangerous fluctuations in river levels.

Our system has great potential to be deployed as a near real-time system for recreational river users. It is particularly successful in forecasting over short forecast windows. Given our ensemble approach, individual weather site models can be trained offline over large historical datasets. In real-time, an interface could perform inference using the ensemble model to provide kayakers with up-to-date future forecasts. We provide an interactive demonstration of our approach online at https://share.streamlit.io/orion-junkins/river-cast-frontend/main/app.py

References

[1] A. Mosavi, P. Ozturk, and K. Chau, “Flood prediction using machine learning models: Literature review,” *Water*, vol. 10, no. 11, p. 1536, 2018.

[2] H. R. Maier, A. Jain, G. C. Dandy, and K. P. Sudheer, “Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions,” *Environmental modelling & software*, vol. 25, no. 8, pp. 891–909, 2010.

[3] W. Li, Q. Duan, C. Miao, A. Ye, W. Gong, and Z. Di, “A review on statistical postprocessing methods for hydrometeorological ensemble forecasting,” *Wiley Interdisciplinary Reviews: Water*, vol. 4, no. 6, p. e1246, 2017.

[4] D. Hussain, T. Hussain, A. A. Khan, S. A. A. Naqvi, and A. Jamil, “A deep learning approach for hydrological time-series prediction: A case study of Gilgit river basin,” *Earth Science Informatics*, vol. 13, no. 3, pp. 915–927, 2020.

[5] W. Xu, Y. Jiang, X. Zhang, Y. Li, R. Zhang, and G. Fu, “Using long short-term memory networks for river flow prediction,” *Hydrology Research*, vol. 51, no. 6, pp. 1358–1376, 2020.

[6] A. M. Ahmed *et al.*, “New double decomposition deep learning methods for river water level forecasting,” *Science of The Total Environment*, vol. 831, p. 154722, 2022.

[7] V. Nourani, “An emotional ANN (EANN) approach to modeling rainfall-runoff process,” *Journal of Hydrology*, vol. 544, pp. 267–277, 2017.

[8] T. C. Pagano *et al.*, “Challenges of operational river forecasting,” *Journal of Hydrometeorology*, vol. 15, no. 4, pp. 1692–1707, 2014.