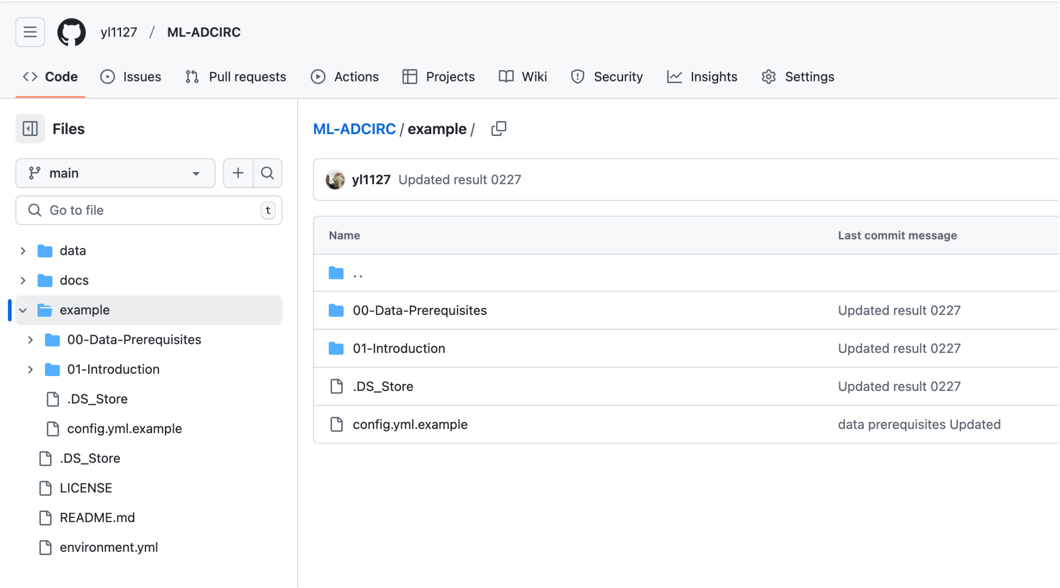
Introduction to Costal Flood Forecasting

Yunlong Pan

Outline

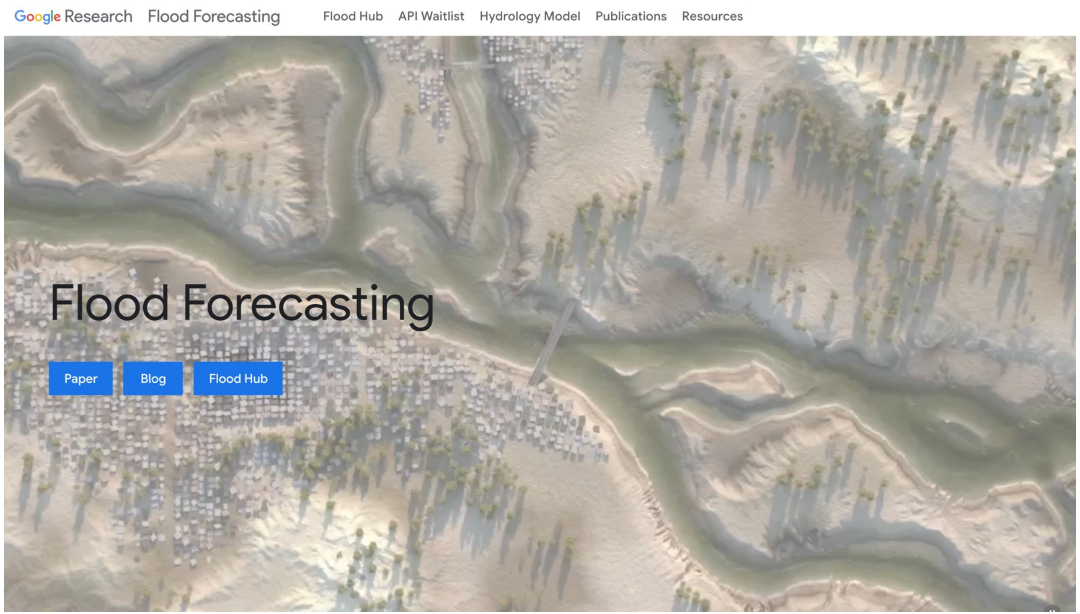
- Introduction
- Initial Result
  - Initial Training
- Method
- Plan
- Discussion



https://github.com/yl1127/ML-ADCIRC

Outline

- Introduction
- Initial Result
  - Initial Training
- Method
- Plan
- Discussion



https://sites.research.google/gr/floodforecasting/

Outline

- Introduction
- Initial Result
  - Initial Training
- Method
- Plan
- Discussion

Article | Open access | Published: 20 March 2024

Global prediction of extreme floods in ungauged watersheds

Grey Nearing, Deborah Cohen, Vusumuzi Dube, Martin Gauch, Oren Gilon, Shaun Harrigan, Avinatan Hassidim, Daniel Klotz, Frederik Kratzert, Asher Metzger, Sella Nevo, Florian Pappenberger, Christel Prudhomme, Guy Shalev, Shlomo Shenzis, Tadele Yednkachw Tekalign, Dana Weitzner & Yossi Matias

Nature **627**, 559–563 (2024) | Cite this article

Abstract

Floods are one of the most common natural disasters, with a disproportionate impact in developing countries that often lack dense streamflow gauge networks[1]. Accurate and timely warnings are critical for mitigating flood risks[2], but hydrological simulation models typically must be calibrated to long data records in each watershed. Here we show that artificial intelligence-based forecasting achieves reliability in predicting extreme riverine events in ungauged watersheds at up to a five-day lead time that is similar to or better than the reliability of nowcasts (zero-day lead time) from a current state-of-the-art global modelling system (the Copernicus Emergency Management Service Global Flood Awareness System). In addition, we achieve accuracies over five-year return period events that are similar to or better than current accuracies over one-year

Download PDF

Associated content

Artificial intelligence can provide accurate forecasts of extreme floods at global scale

Nature | **Research Briefing** | 03 Apr 2024

Sections | Figures | References

https://www.nature.com/articles/s41586-024-07145-1

Introduction

- Flooding type
  - River flood
  - Coastal flood
  - Storm surge
  - Flash flood

SEVERE WEATHER 101

Flood Types

A **river flood** occurs when water levels rise over the top of river banks due to excessive rain from tropical systems making landfall, persistent thunderstorms over the same area for extended periods of time, combined rainfall and snowmelt, or an ice jam.

River flooding occurs when water levels rise over the top of river banks due to excessive rain from tropical systems making landfall, persistent thunderstorms over the same area for extended periods of time, combined rainfall and snowmelt, or an ice jam. [+]

A **coastal flood**, or the inundation of land areas along the coast, is caused by higher than average high tide and worsened by heavy rainfall and onshore winds (i.e., wind blowing landward from the ocean). Places like Charleston, South Carolina, and Savannah, Georgia, experience impacts from shallow coastal flooding several times a year because of coastal development and lower elevation.

Storm surge is an abnormal rise in water level in coastal areas, over and above the regular astronomical tide, caused by forces generated from a severe storm's wind, waves, and low

EDUCATION

- Severe Weather 101
- For Educators
- For Students
- For Everyone

SEVERE WEATHER 101

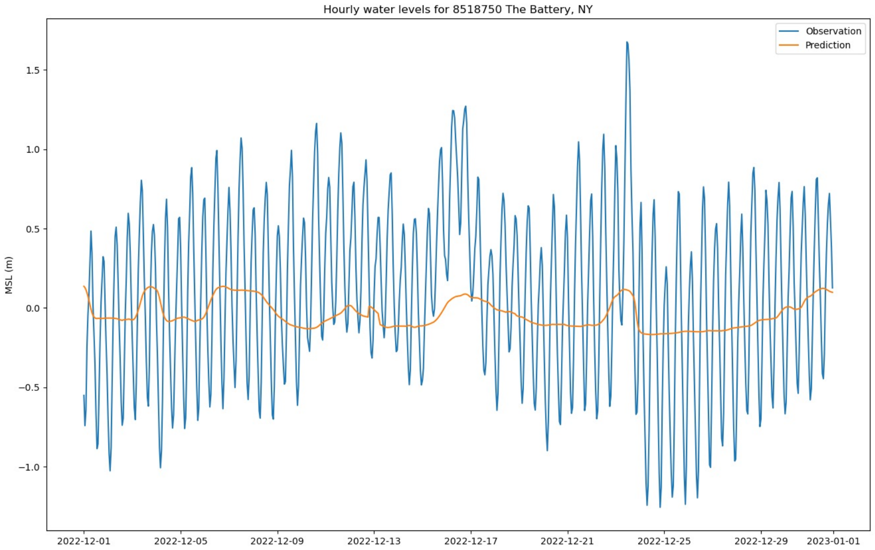
- Thunderstorms
- Tornadoes
- Floods
  - Basics
  - Types
  - Detection
  - Forecasting
  - FAQ
- Lightning
- Hail

https://www.nssl.noaa.gov/education/svrwx101/floods/types/

Initial Result

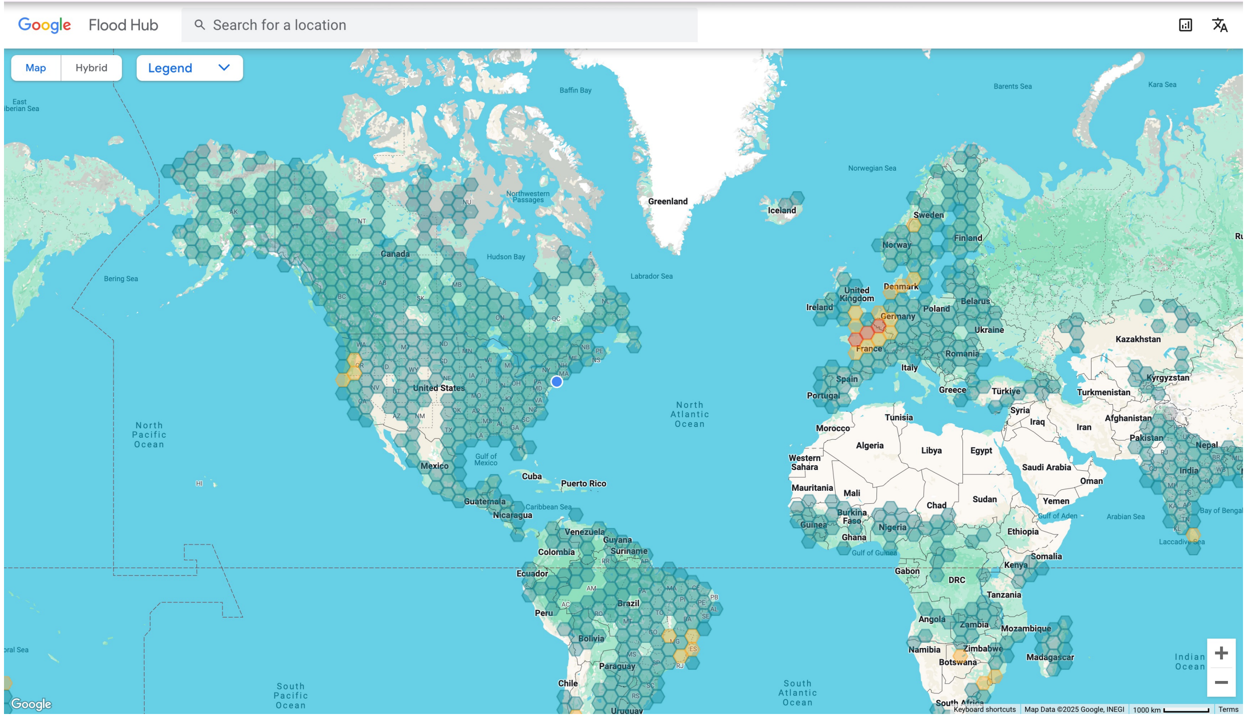
- Location: The Battery, NY
- Station ID: 8518750
- Time: 2022-12-01 to 2022-12-31(hourly)

- It works. But not good yet.



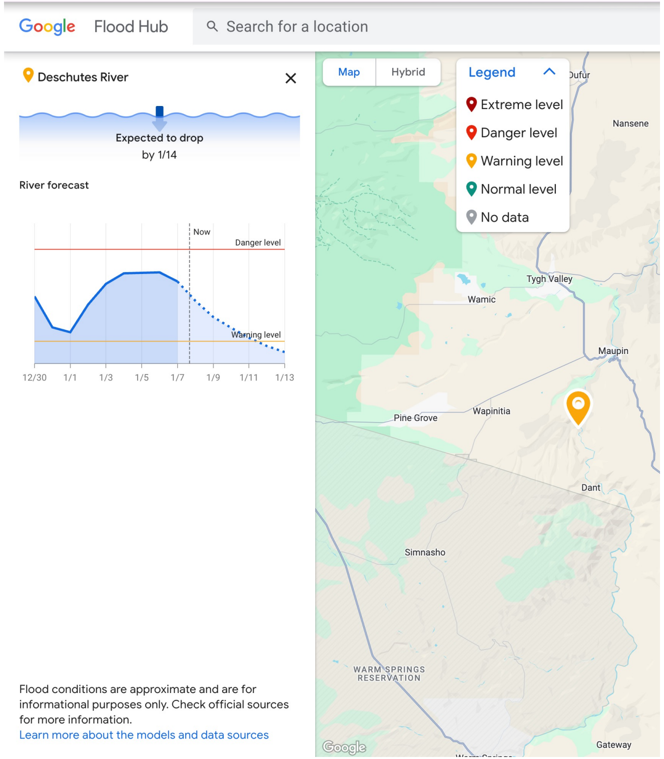
https://github.com/yl1127/ML-ADCIRC/blob/main/example/01-Introduction/Introduction.ipynb

Product: Flood Hub



https://sites.research.google/floods/

Return Period



Models and data sources

Availability

In which regions is Flood Hub available?

Our models are currently covering more than 5000 locations across river basins in 101 countries. The expert mode allows showing information in over 150 countries, and over 245k locations. Our research teams are working tirelessly to develop ways to expand our coverage.

What information is provided in flood forecasting?

1. A map of current and expected floods.
2. River changes forecasted over time, including alert thresholds that represent 2, 5 and 20 years return period.
3. In some cases, depending on data availability, an illustration of water depth compared to the human body, e.g. ankle height or waist height.
4. Gauge information - see in expert mode when selecting a gauge.

Benchmark (River paper)

- Flooding type
- Data
  - Six variables
- Benchmark
  - GloFAS

  (Global Flood Awareness System)

  the same six variables (ERA5)
- Model

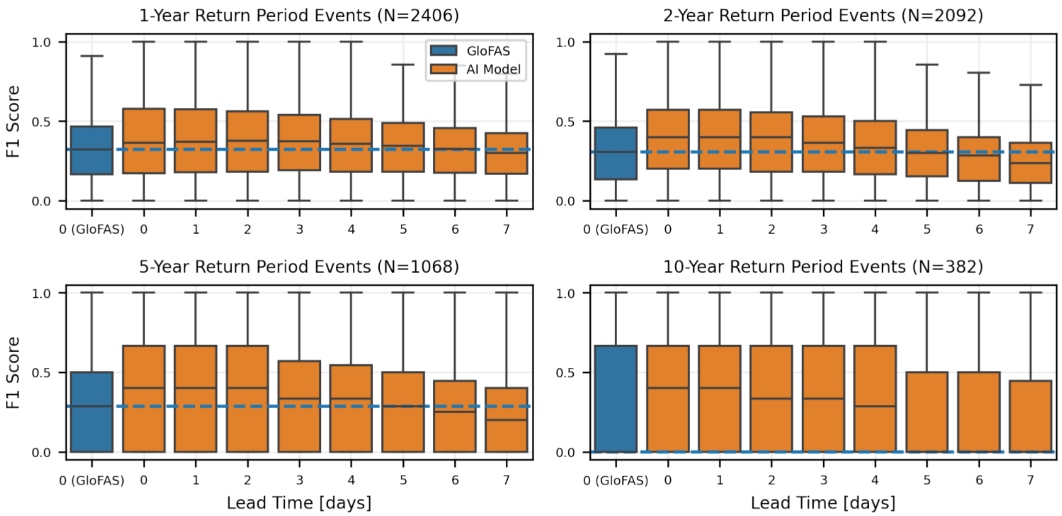


Fig. 3 Distributions over F1 scores at all evaluation gauges as a function of lead time for different return periods. The AI model had similar (not statistically different) or better reliability over 1, 2, and 5 year return periods at 5-day lead time than GloFAS at 0-day lead time. Statistical tests are reported in the main text. Boxes show distribution quartiles and whiskers show the full range excluding outliers. The blue dashed line is the median score for GloFAS nowcasts, and is plotted as a reference.

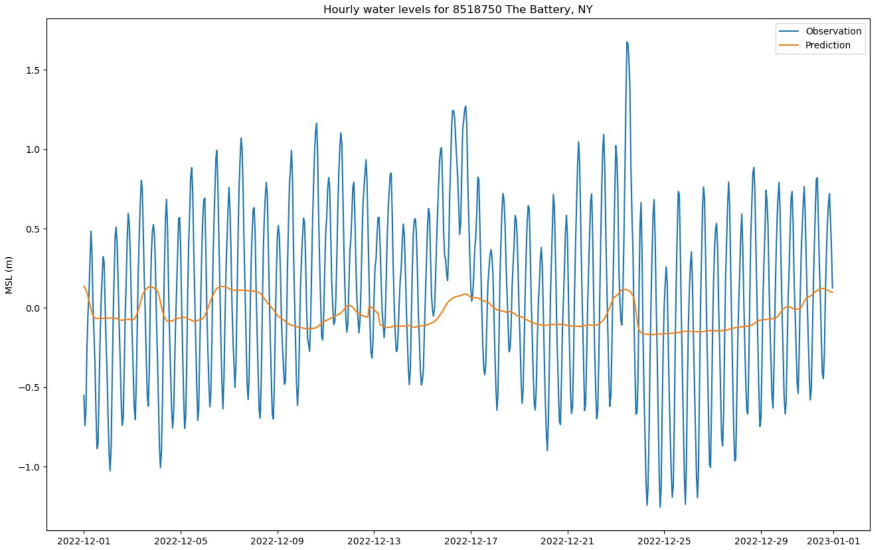
Initial Result

- Location: The Battery, NY
- Station ID: 8518750
- Time: 2022-12-01 to 2022-12-31(hourly)

```
values = metrics.calculate_all_metrics
for key, val in values.items():
    print(f"{key}: {val:.3f}")
```

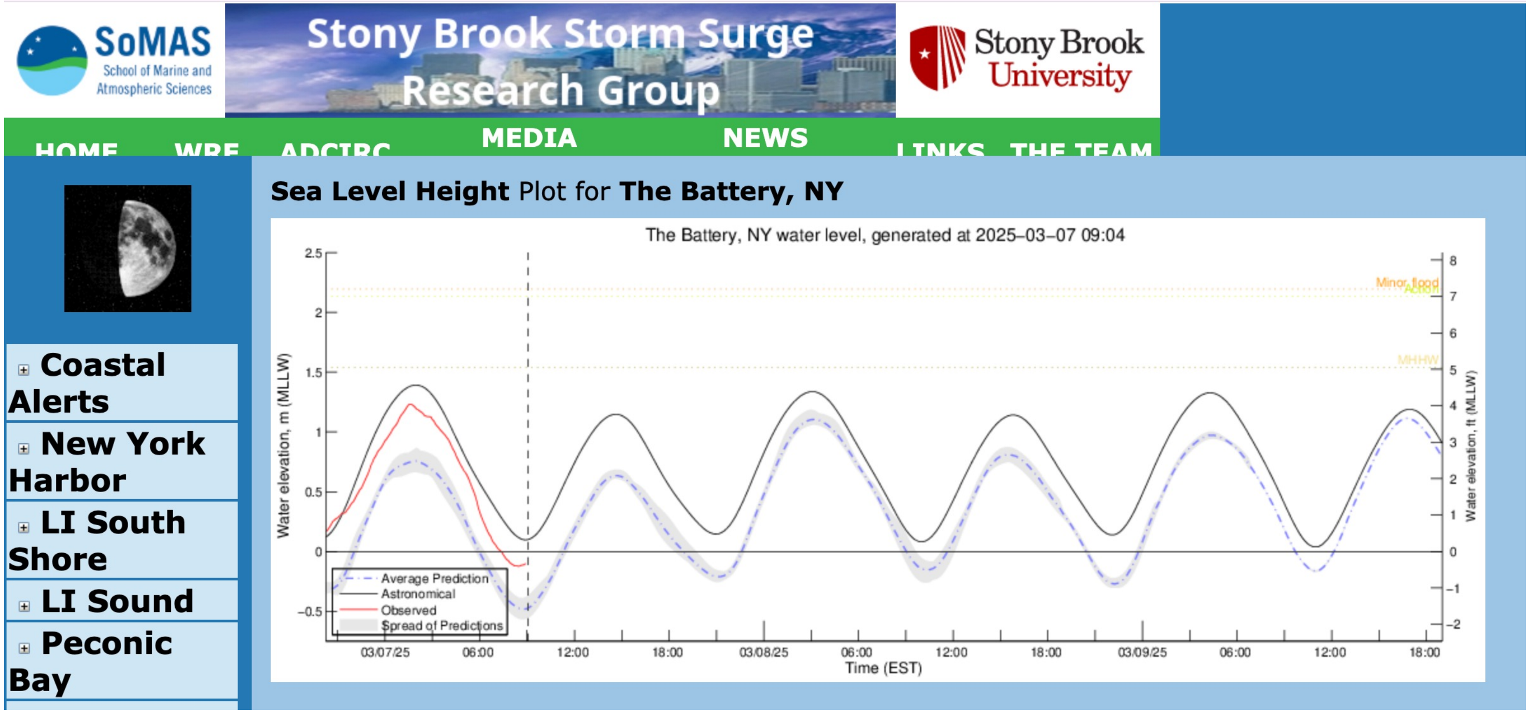
[19]

```
NSE: -0.021
MSE: 0.299
RMSE: 0.547
KGE: -0.825
Alpha-NSE: 0.161
Beta-KGE: -0.423
Beta-NSE: -0.259
Pearson-r: 0.222
FHV: -89.778
FMS: nan
FLV: nan
Peak-Timing: nan
Peak-MAPE: 103.061
```



https://github.com/yl1127/ML-ADCIRC/blob/main/example/01-Introduction/Introduction.ipynb

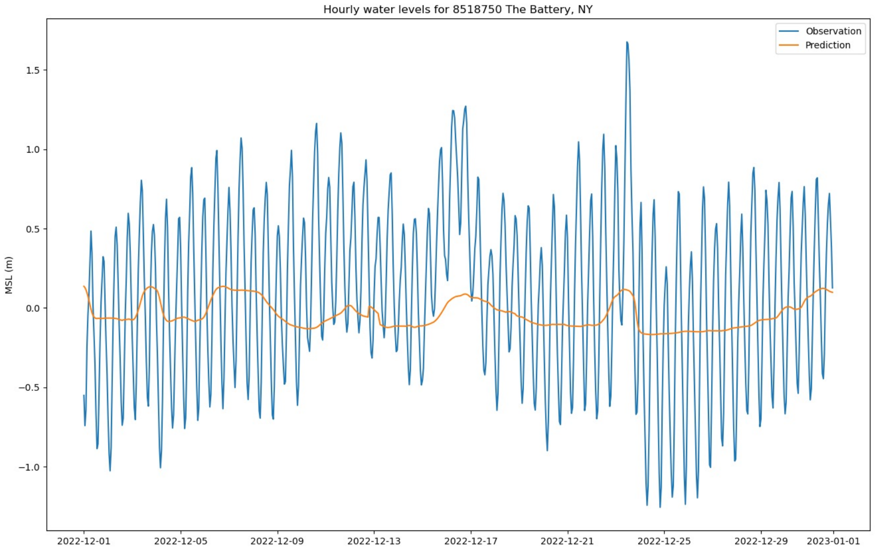
Initial target



https://stormy.msrc.sunysb.edu/

Initial Training

- Dataset: CORA (Water level)
- Dataset: ERA5 (temperature, humidity)
- Location: The Battery, NY
- Station ID: 8518750
- Time:
  - 20220101-20221031 as training
  - 20221101-20221130 as developing
  - 20221201-20221231 as testing
- Model: CudaLSTM from Neuralhydrology
- More details: https://github.com/yl1127/ML-ADCIRC/blob/main/example/01-Introduction/



https://github.com/yl1127/ML-ADCIRC/blob/main/example/01-Introduction/

Coastal Ocean Reanalysis (CORA)

- Location: Gulf of Mexico, Atlantic, and Caribbean
- Time range: 1979-2022
- Data: Hourly waves and water levels
- Methods: ADCIRC+SWAN

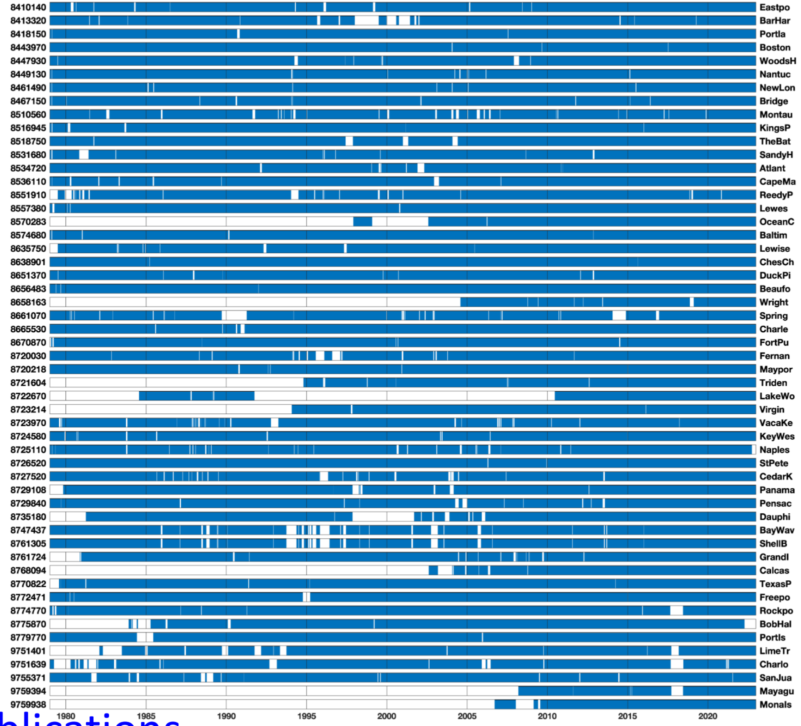
CORA helps create a more complete and consistent picture of historical water levels by modeling waves and water levels between NOAA tide gauges. CORA pairs historical observations from NOAA tide gauges with modern computer models to fill gaps in historical records.

This reanalysis is made possible through observations from National Water Level Observation Network (NWLON). This celebrated, novel approach of combining models and observations can be used to better assess long-term sea level change and compare current flood risks to those of the past, especially in areas where such data is currently unavailable.

https://tidesandcurrents.noaa.gov/cora.html

NOAA vs CORA

- Location: Gulf of Mexico, Atlantic, and Caribbean
- Time range: 1979-2022
- Data: Hourly waves and water levels
- Methods: ADCIRC+SWAN



https://tidesandcurrents.noaa.gov/cora.html#publications

Figure 8. A plot of water level data availability for 53 stations used in the reanalysis process over the period of 1979-2022, with station IDs on the left y-axis and a station name abbreviation on the right. Dates from 1980-2020 are listed in 5-year increments with blue lines delineating completeness of each station's water level time series. Stations are listed from Maine to Texas to Puerto Rico (last 5).

Model (River paper)

- Model
  - LSTM
- Configurations

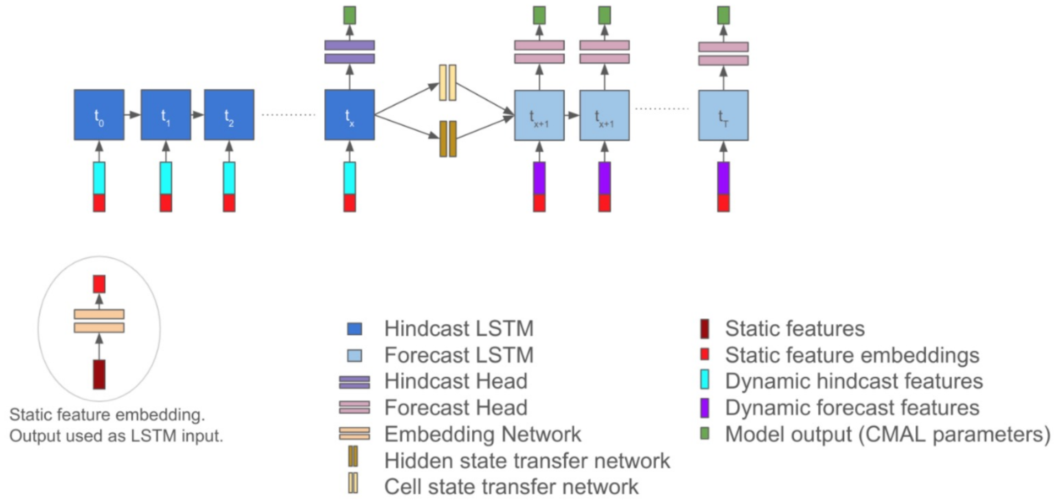
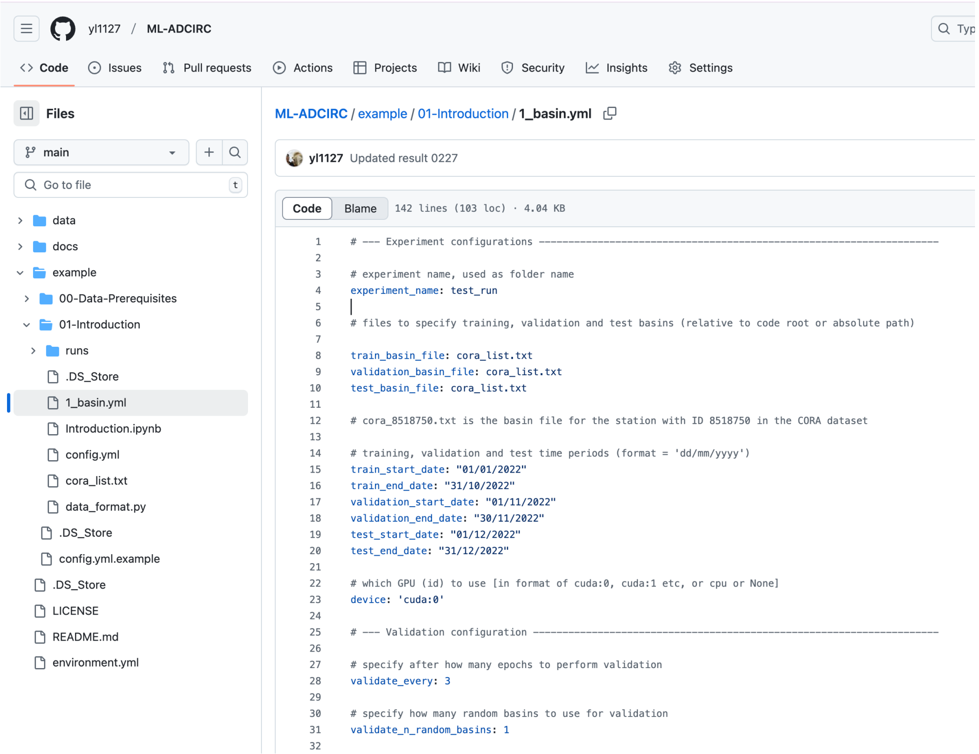


Fig. 9 Architecture of the LSTM-based forecast model developed for this project. This is the model used operationally to support the Google Flood Hub https://g.co/floodhub.

https://neuralhydrology.github.io/

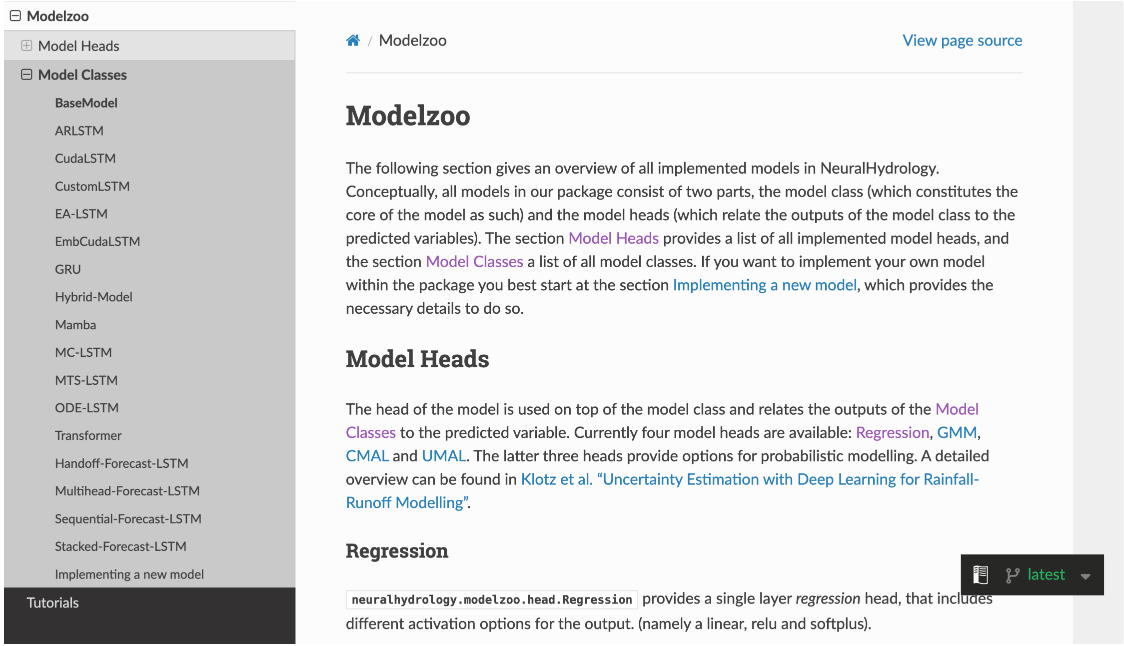
Model (River paper)

- Model
- Configurations
  - model: ealstm
  - hidden_size: 256
  - batch_size: 256
  - epochs: 50,000
  - predict_last_n: 7
  - seq_length: 365



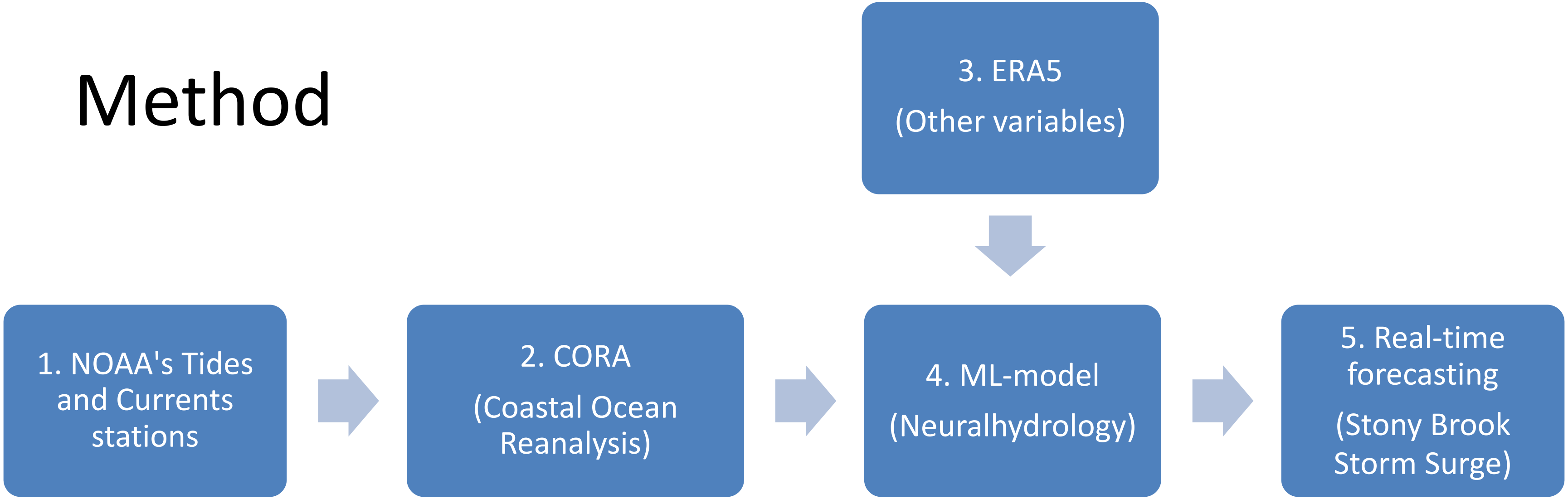
Model

- Model
  - NeuralHydrology
  - LSTM



https://neuralhydrology.readthedocs.io/en/latest/usage/models.html#basemodel

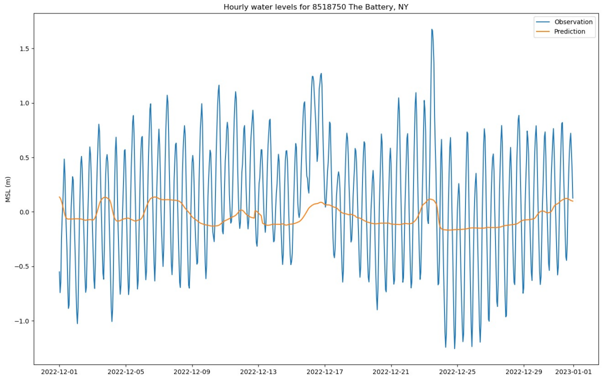
Method

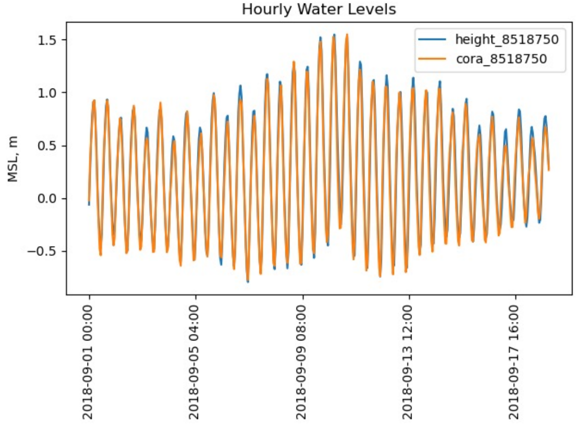


- https://tidesandcurrents.noaa.gov/stations.html?type=Water+Levels
- https://tidesandcurrents.noaa.gov/cora.html
- https://github.com/google-research/arco-era5
- https://neuralhydrology.readthedocs.io/en/latest/index.html
- https://stormy.msrc.sunysb.edu/

Plan

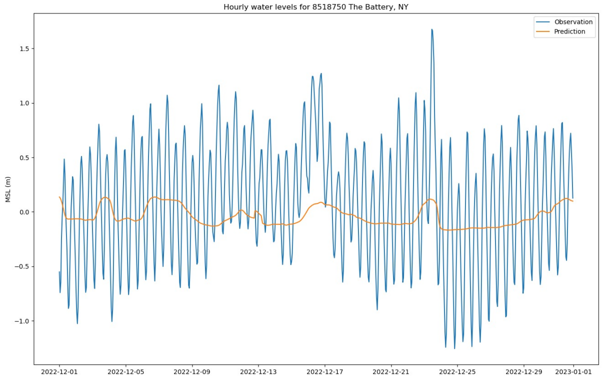
- Train a better model
  - Location: The Battery, NY - Station ID: 8518750
  - Data
    - Time: 1979-01-01 to 2022-12-31
    - How to choose the weather variables
  - Model
    - How to select the best model?
    - How to do the config.yml?
  - Benchmark
    - How's the Stony Brook Storm Surge result?

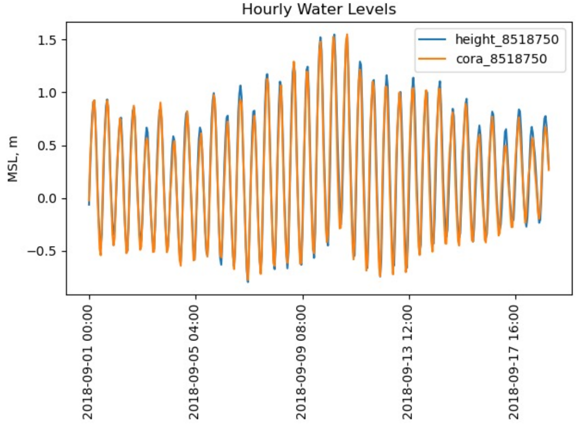




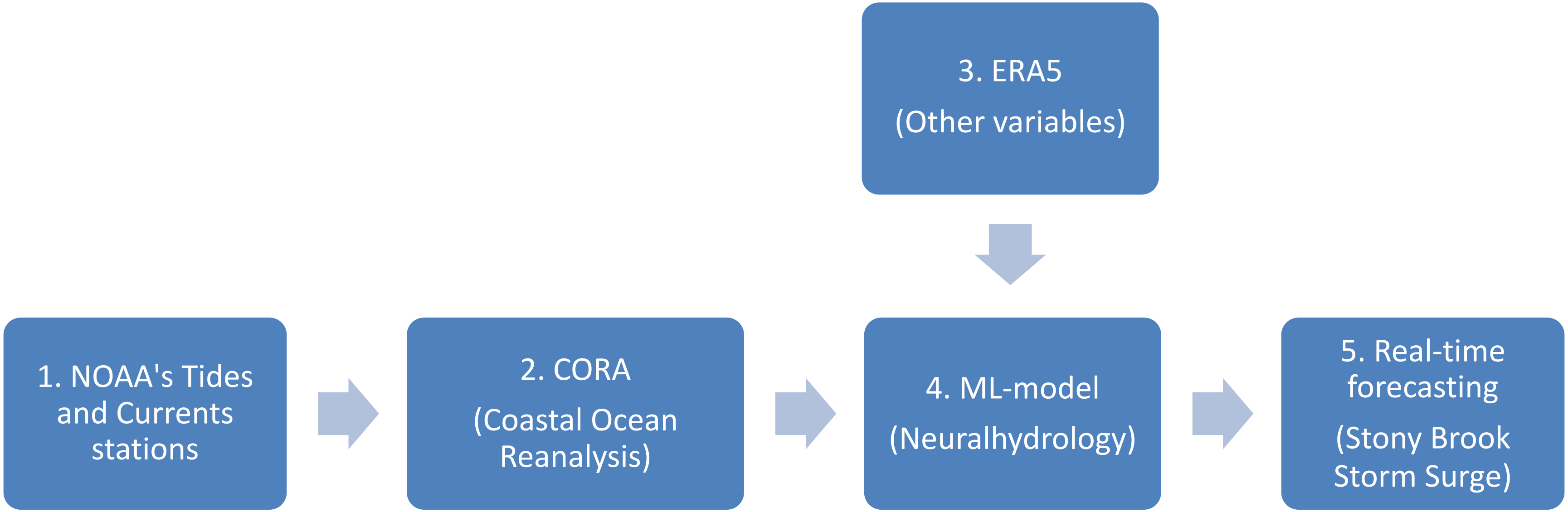
Plan

- Long-term:
  - More stations
  - Real-time
  - Learn CORA dataset in deep
  - New Methods
  - Product (e.g. Flood hub)





Discussion


3. ERA5
(Other variables)
1. NOAA's Tides and Currents stations
2. CORA
(Coastal Ocean Reanalysis)
4. ML-model
(Neuralhydrology)
5. Real-time forecasting
(Stony Brook Storm Surge)