Group Meeting

Yunlong Pan

Outline

- GitHub
  - https://github.com/yl1127/ML-ADCIRC
- Dataset: Coastal Ocean Reanalysis (CORA)
- Flood Forecasting
  - Flooding type
  - Data
  - Benchmark
  - Model

Coastal Ocean Reanalysis (CORA)

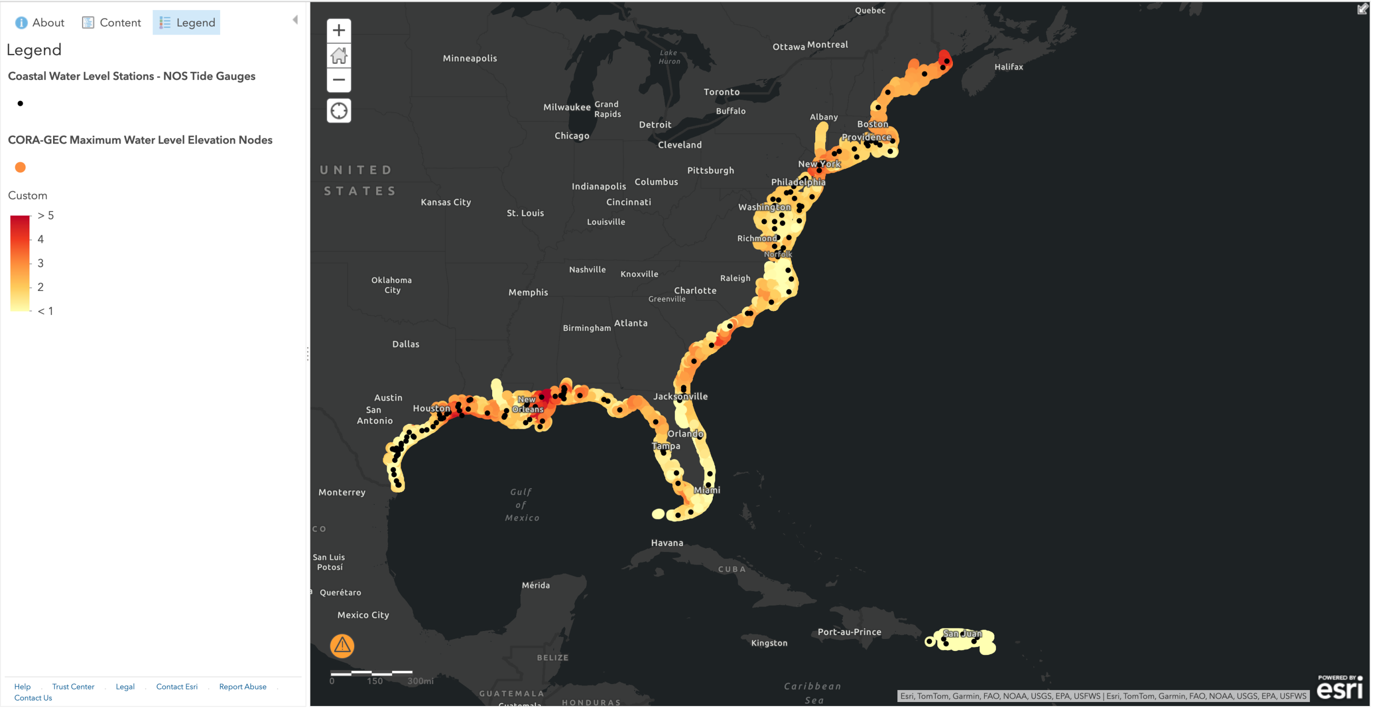
- Loction: Gulf of Mexico, Atlantic, and Caribbean
- Time range: 1979-2022
- Data: Hourly waves and water levels
- Methods: ADCIRC+SWAN

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

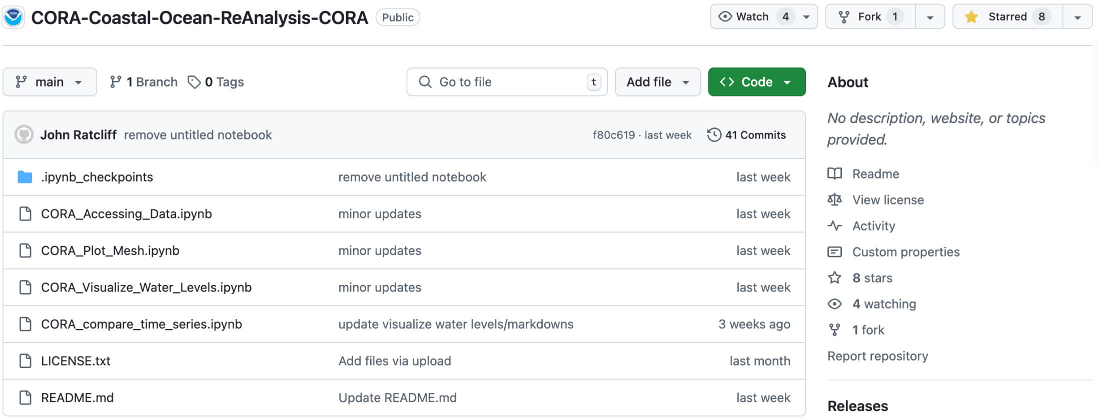
CORA: Example



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

CORA: Strengthen

- 1. It's exactly what we want, and even more
- 2. It's high quality and high resolution
- 3. It's new



https://github.com/NOAA-CO-OPS/CORA-Coastal-Ocean-ReAnalysis-CORA

Flood Forecasting

- Flooding type
  - River flood
- Data
  - Weather data
  - Geological data
  - Streamflow
- Benchmark
  - GloFAS
- Model
  - LSTM

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

Article

Global prediction of extreme floods in ungauged watersheds

https://doi.org/10.1038/s41586-024-07145-1

Received: 29 July 2023

Accepted: 31 January 2024

Published online: 20 March 2024

Open access

Check for updates

Grey Nearing[1✉], Deborah Cohen[1], Vusumuzi Dube[1], Martin Gauch[1], Oren Gilon[1], Shaun Harrigan[2], Avinatan Hassidim[1], Daniel Klotz[3], Frederik Kratzert[1], Asher Metzger[1], Sella Nevo[4], Florian Pappenberger[2], Christel Prudhomme[2], Guy Shalev[1], Shlomo Shenzis[1], Tadele Yednkachw Tekalign[1], Dana Weitzner[1] & Yossi Matias[1]

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 return period events. This means that artificial intelligence can provide flood warnings earlier and over larger and more impactful events in ungauged basins. The model developed here was incorporated into an operational early warning system that produces publicly available (free and open) forecasts in real time in over 80 countries. This work highlights a need for increasing the availability of hydrological data to continue to improve global access to reliable flood warnings.

Floods are the most common type of natural disaster[3] and the rate of flood-related disasters has more than doubled since 2000[4]. This increase in flood-related disasters is driven by an accelerating hydrological cycle caused by anthropogenic climate change[5,6]. Early warning systems are an effective way to mitigate flood risks, reducing flood-related fatalities by up to 43%[7,8] and economic costs by 35–50%[9,10]. Populations in low- and middle-income countries make up almost 90% of the 1.8 billion people that are vulnerable to flood risks[1]. The World Bank has estimated that upgrading flood early warning systems in developing countries to the standards of developed countries would save an average of 23,000 lives per year[2].

In this paper, we evaluate the extent to which artificial intelligence (AI) trained on open, public datasets can be used to improve global access to forecasts of extreme events in global rivers. On the basis of the model and experiments described in this paper, we developed an operational system that produces short-term (7-day) flood forecasts in over 80 countries. These forecasts are available in real time without barriers to access such as monetary charge or website registration (https://g.co/floodhub).

A major challenge for riverine forecasting is that hydrological prediction models must be calibrated to individual watersheds using long data records[11,12]. Watersheds that lack stream gauges to supply data for calibration are called ungauged basins, and the problem of 'prediction in ungauged basins' (PUB) was the decadal problem of the International Association of Hydrological Sciences (IAHS) from 2003 to 2012[13]. At the end of the PUB decade, the IAHS reported that little progress had been made against the problem, stating that "much of the success so far has been in gauged rather than in ungauged basins, which has negative effects in particular for developing countries"[14].

Only a few per cent of the world's watersheds are gauged, and stream gauges are not distributed uniformly across the world. There is a strong correlation between national gross domestic product and the total publicly available streamflow observation data record in a given country (Extended Data Fig. 1 shows this log–log correlation), which means that high-quality forecasts are especially challenging in areas that are most vulnerable to the human impacts of flooding.

In previous work[15], we showed that machine learning can be used to develop hydrological simulation models that are transferable to ungauged basins. Here we develop that into a global-scale forecasting system with the goal of understanding scalability and reliability. In this paper, we address whether, given the publicly available global streamflow data record, it is possible to provide accurate river forecasts across large scales, especially of extreme events, and how this compares with the current state of the art.

The current state of the art for real-time, global-scale hydrological prediction is the Global Flood Awareness System (GloFAS)[16,17]. GloFAS is the global flood forecasting system of Copernicus Emergency Management Service (CEMS), delivered under the responsibility of the European Commission's Joint Research Centre and operated by the European Centre for Medium-Range Weather Forecasts (ECMWF) in

Flood Forecasting

- Flooding type
  - Coastal flood (Target Variables: water level)
  - Storm surge (Target Variables: water level)
- Data
  - Coastal Ocean Reanalysis (CORA)
- Benchmark (To do)
  - NOAA Tide Predictions https://tidesandcurrents.noaa.gov/tide_predictions.html
  - NOAA Water Levels https://tidesandcurrents.noaa.gov/stations.html?type=Water+Levels
  - NOAA 1-Minute Water Level Data https://tidesandcurrents.noaa.gov/1mindata.html
  - NOAA Extreme Water Levels https://tidesandcurrents.noaa.gov/est/
  - NOAA Coastal Inundation Dashboard https://tidesandcurrents.noaa.gov/inundationdb_info.html
  - SOMAS Stony Brook Storm Surge Research Group https://stormy.msrc.sunysb.edu/
- Model

Review: Result

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

  (Global Flood Awareness System)

  the same six variables (ERA5)
- Model

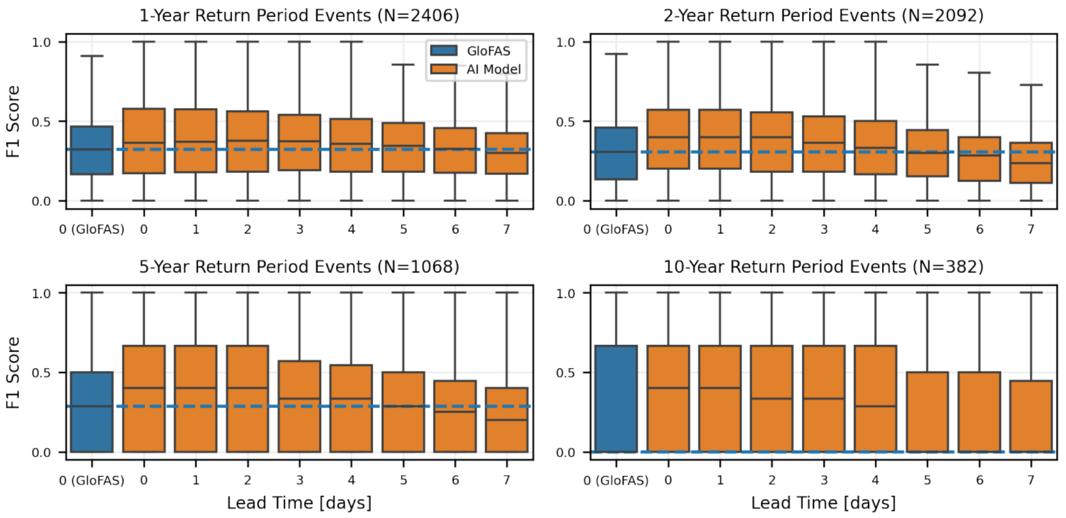


Fig. 3 Distributions over F1 scores at all evaluation gauges as a function of lead time for different return periods. The AI model had similar (not statistically different) or better reliability over 1, 2, and 5 year return periods at 5-day lead time than GloFAS at 0-day lead time. Statistical tests 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.

Discussion