

The Google Flood Forecasting Initiative

Sella Nevo¹[0000-0002-4743-3634]

Google Research, Tel Aviv

Abstract. The Google Flood Forecasting Initiative is the world’s first large-scale machine-learning-based operational flood forecasting system. It currently covers more than 360 million people, and provides flood warnings to governments, disaster management organizations, and individuals at risk. It combines two separate models - a hydrologic model and an inundation model - which together can provide high accuracy, actionable forecasts for flooding in upcoming days.

Keywords: Flood forecasting · Hydrology · LSTM.

1 Introduction

Floods are responsible for 5,000 to 60,000 fatalities [5, 18, 4, 17, 12, 11], affect between 95 million and 250 million people [5, 13, 12, 17, 4, 22], and cause between \$21 and \$33 billion in economic damages annually [3–5]. The majority of these harms are caused by riverine floods [17] - where a river overflows its banks and inundates the floodplain around it. The frequency and severity of floods has been increasing in recent decades [21, 10], and are expected to rise significantly in the future due to climate change, land use changes and other long term processes [16, 15]. On a brighter note, reliable early warning systems have been shown to prevent up to 43% of fatalities [2, 26], and up to 50% of economic damages [23, 1]. They have also been shown to have an impressive cost-benefit ratio of 1:9 [6], and much higher in low and middle-income countries (LMICs) [25], making them a top climate change adaptation policy recommendation by the United Nations and the World Bank.

Human costs of flooding are heavily concentrated in LMICs - for example, about half of all flood-related deaths globally occur in India and Bangladesh [17]. Despite the promise of early warning systems, many of these countries lack the resources for the comprehensive data collection, ongoing recalibration and maintenance, and computational resources that classic flood forecasting systems demand.

The Google Flood Forecasting Initiative aims to provide accurate and actionable flood warnings globally. It utilizes machine learning to create forecasting systems that are both more accurate and more scalable than classic flood forecasting systems, and then disseminates the resulting warnings via Google’s various interfaces (Search, Maps, and smartphone notifications) as well as specialized interfaces to support governments and NGOs engaged in disaster management. At the time of writing, the system provides flood forecasting systems to regions covering more than 360 million people.

2 Problem Statement

Flood forecasting systems aim to provide governments, NGOs and individuals in the affected region reliable and actionable information about upcoming floods. The modeling task of riverine flood forecasting systems can often be divided into two main sub-tasks.

The first is forecasting conditions within the river, sometimes described as hydrologic modeling. In this task one traditionally uses inputs such as precipitation, temperature, soil moisture and others throughout the river basin, and produces as an output either the river discharge or the water level at specific points within the river at a specific time, usually up to two weeks into the future. Traditional hydrologic models are “conceptual” models - they are inspired by the actual physical processes involved, though usually are vastly simpler than the actual processes they describe. They will have somewhere between several to several dozen parameters, which are calibrated to produce outputs that fit the historical measurements in the specific river.

The second task is forecasting behavior across the floodplain, sometimes described as inundation modeling. Here we assume we already know the discharge or water level in the river (provided by a hydrologic model or a real-time measurement), and model the movement of water across the floodplain. The goal is to produce a spatially accurate map of flood extent (which areas are flooded) or flood depth (how deep is the water in each point on the map). The most common practice for these types of models are called hydraulic models - these are physics-based models which find finite-element solutions to a set of differential equations (specifically the St. Venant equations [8]) on a grid. These models can describe water behavior across the floodplain very accurately, when all their input data is accurate and they are run at sufficient resolution.

The case for incorporating AI into these two types of models is not identical, but similar principles apply. Both traditional models (hydrologic and hydraulic models) require significant manual calibration, and often continuous recalibration based on new information. This leads to high costs of deployment and operation, up to hundreds of thousands of dollars for a single basin - which significantly limits their availability in the regions that need warning systems the most, and prevents scale-up of high-quality systems to truly global scales. Relatedly, both have extremely limited spatial transferability - researchers have been largely unsuccessful in utilizing models trained in one location to improve performance in other locations, e.g. ones with less data available. Computational costs can also act as a critical limitation - especially for the hydraulic model, which is incredibly sensitive to resolution yet requires computation proportional to one over the resolution cubed. Finally, both models share a limitation that is common in physics-based and conceptual models - they can be overly rigid. Even when repeated historical data shows some assumption of the model to be incorrect, it is often difficult (and sometimes impossible) to change the model to correct its errors.

3 Method

Our flood forecasting system uses ML-based methods to tackle each of the tasks described above. We will describe our models for each task separately. For each task, we currently have two different architectures in production - used in different circumstances, e.g. depending on data availability and performance. See figure 1 for an overview of the full system including all components described below, and other components beyond the scope of this chapter.

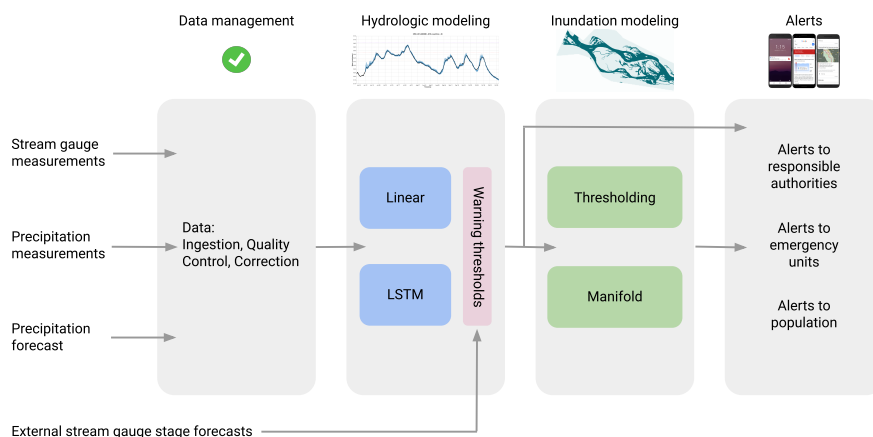


Fig. 1. A high-level overview of the Google flood forecasting system.

3.1 Hydrologic Model

The two architectures used for hydrologic modeling are linear models, and LSTMs. All our models currently forecast water levels (as opposed to discharge), and provide forecasts in gauged locations (i.e. locations where there is a stream gauge installed) - though work is ongoing to expand our reach beyond this. The training and validation scheme for both models is described in a sub-section after the models themselves.

Linear Model The model input includes past river stages from the target gauge (the gauge being forecast) and its upstream gauges (typically 2-5 gauges) for each time step (usually hourly). The output is a future river stage at the target gauge for a given lead time. A multiple linear regression model is trained with historical records of the above inputs and outputs. The model is optimized

using the mean square error (MSE) loss function with L2 regularization. Linear models are trained separately for each target gauge, with a separate linear model trained for all lead times up to the gauge’s maximal lead time. For example, a target gauge with a selected maximal lead time of 24 hours and hourly resolution implies 24 trained linear models. Figure 2 illustrates the schema of the linear model.

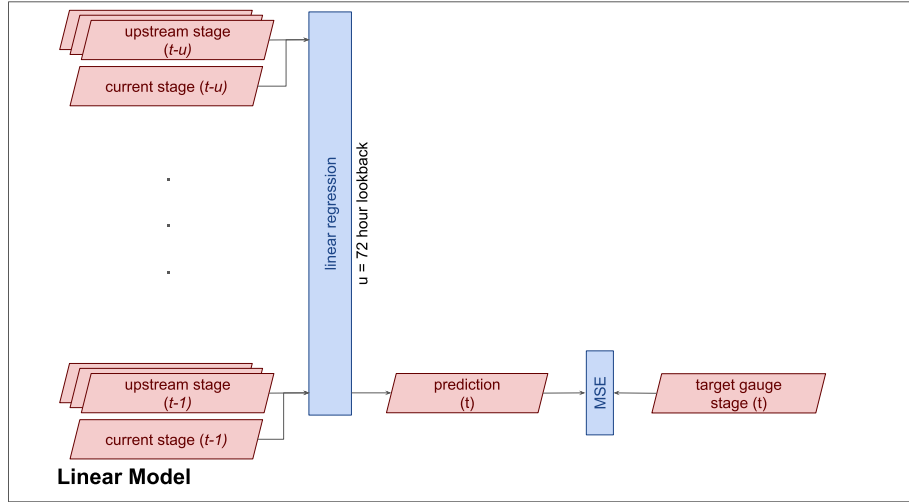


Fig. 2. Schema of the linear model.

Long Short-Term Memory Network (LSTM) Model Building on the work described in [19, 20, 14], the LSTM model consists of two LSTMs: a sequence-to-one hindcast model and a sequence-to-sequence forecast model, where the output of the hindcast model is used as the initial state of the forecast model. The hindcast model processes data from past days sequentially, taking the following variables as inputs at each timestep: (i) Precipitation measurements through the basin, (ii) near-real-time stage, and (iii) a linear combination of stage measurements from upstream gauges over some upstream “lookback” period (typically a few days). The linear combination of stage measurements is produced by a separate linear layer with gauge-specific weights that combines a (variable) number of upstream inputs per gauge into five features that are fed as inputs into the hindcast LSTM (similar to the linear model described above).

The hindcast LSTM runs until the current time, defined to be the time of the last available measurement. The final cell state and hidden state of the hindcast LSTM are passed through a fully connected layer and the output of this “state handoff” layer is used as the initial cell state and hidden state of the forecast

LSTM. The forecast LSTM advances one step for every lead time, producing the relevant forecasts for all lead times up to the maximal one. The main goal of using a state handoff between two LSTMs is to distinguish between the different inputs that are available in real-time during hindcast and forecast - and specifically the availability of precipitation measurements.

All weights of the LSTMs (hindcast and forecast) are shared between all target gauges, i.e. they are regionally calibrated. The only gauge-specific weights are those of the linear upstream combiner layer which allows the architecture to support varying numbers of upstream gauges, travel times, sizes of tributaries, etc.

The system estimates the uncertainty of the water stage, or the time dependent distribution over the predicted stage, using a countable mixture of asymmetric Laplacians (CMAL). The parameters of this distribution are generated by feeding the hidden state of the forecast LSTM into a dedicated head layer for each forecasted time step. At each time step, the loss is calculated as the negative log-likelihood of the observed stages given the LSTM forecasts using the next maximal lead time values. It is important to note that the likelihood-based loss function is calculated only over the outputs of the forecast LSTM, and the hindcast LSTM is used only to initialize the forecast state. Since training is shared for all target gauges, the maximal lead time in the training phase is taken as the maximum of the gauge-specific values. Figure 3 illustrates the schema of the LSTM model.

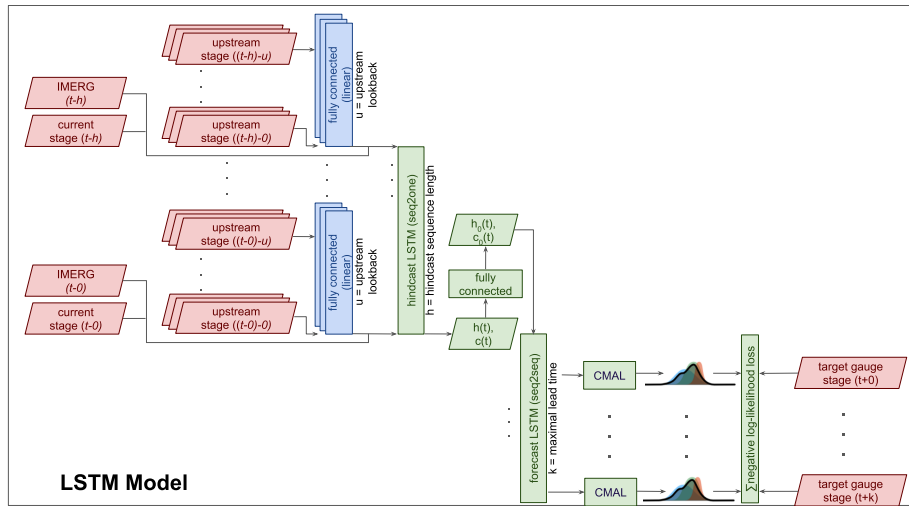


Fig. 3. Schema of the LSTM model.

Training and Validation The stage forecast models are trained and validated with historical data using a cross-validation scheme, in which each fold uses one year’s worth of data for validation and the rest for training (i.e., 1-year leave-out). For operational use, the models were retrained on the full data set to produce the best real-time forecasts possible.

Operational Use For Alerting For each target gauge, if the maximal forecasted river stage between the forecast’s “current time” and the gauge’s maximal lead time is above the predefined gauge-specific warning threshold, an alert is issued, and this maximal stage is used for inundation mapping (see below).

3.2 Inundation Model

The two architectures used for inundation modeling are the Thresholding model, and the Manifold model, both described below. The Thresholding model produces an inundation extent map but no forecast for flood depths. The Manifold model additionally produces the water depth at each pixel, but requires a digital elevation model (DEM) and requires more effort to implement and deploy at scale. The training and validation scheme for both models is described in a sub-section after the models themselves.

Thresholding Model The model assumes that each pixel in the area of interest (AOI) becomes inundated when the target gauge exceeds a (pixel-specific) threshold water stage. These thresholds are learned from the series of historic stage data at the target gauge and the corresponding state of the pixel (dry/wet) during these events. Each pixel in the inundation map is treated as a separate classification task, predicting whether the pixel will be inundated or not. We refer to the “wet” class as the positive class.

The algorithm described below identifies pixel-specific thresholds and is aimed at maximizing some F-score using an optimized global parameter called minimal ratio. An F-score [7] refers to a weighted geometric mean of precision and recall, with F_1 referring to the simple harmonic mean between them and F_3 giving more weight to precision over recall, for example. The algorithm below can be optimized for any choice of F-score. To achieve this, an iterative process is applied to each pixel. In each iteration, we find the threshold that maximizes the ratio of true wet events (where the water stage at the gauge is above the threshold and the pixel was wet) to false wet events (where the water stage at the gauge is above the threshold and the pixel is dry). The threshold that maximizes this ratio is the most cost-effective threshold in the sense that it provides the most true wets per false wet instance. At the first iteration all training events are considered; then, after each selection of a threshold and its respective true-false ratio, events with stage measurements above the threshold are discarded and a new iteration starts with the remaining events. If the new true-false ratio calculated is lower than the minimal ratio parameter value, the process stops and the final threshold for the pixel is the one found in the previous iteration.

It can be shown that for every minimal ratio parameter value, no other set of pixel-specific thresholds achieves simultaneously better precision (i.e., fraction of all flooded pixels that are predicted as being flooded) and recall (i.e., fraction of all pixels that are predicted to be flooded and are really flooded); implying it is Pareto optimal. Therefore, for any F-score there exists some value of the minimal threshold parameter which finds the thresholds that optimize this F-score.

In cases where the river stage input is higher than all past stage data, the Thresholding model’s output inundation map is initialized from the most severe inundation extent seen in the historical events and expanded in all directions. The expansion distance is a linear function of the difference between the forecasted stage and the stage of the highest historical event.

This Thresholding model requires almost no site-specific data like DEMs, and no manual work, making it appealing for large-scale deployment across many areas of interest in a short amount of time.

Figure 4 illustrates the schema of the Thresholding model.

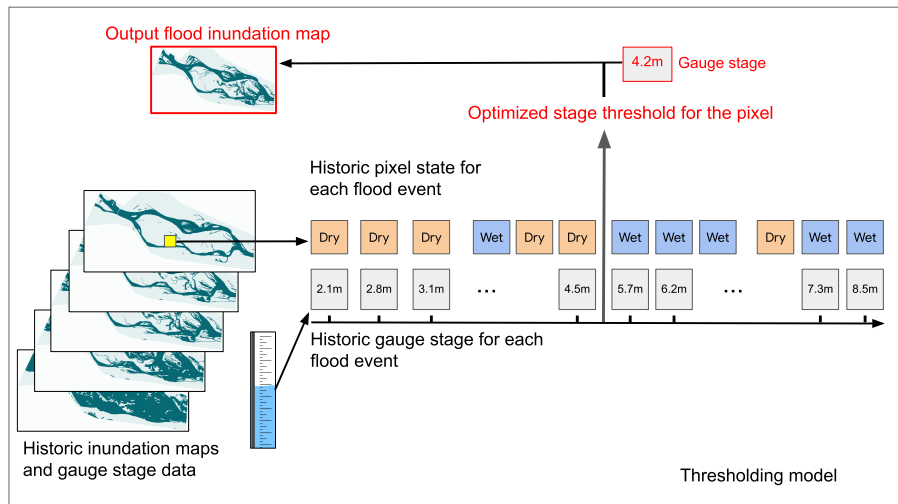


Fig. 4. Schema of the Thresholding model.

Manifold Model The Manifold inundation model provides a machine-learning alternative to hydraulic models, by computing physically reasonable flood inundation. Its inputs are a DEM for the AOI and a target water stage. It outputs both the flood inundation extent and the inundation depth at each AOI pixel. The model is divided into two major parts, described below.

Flood Extent To Water Height Algorithm

The flood extent to water height algorithm converts a DEM and an inundation

extent map (i.e., wet/dry state for each pixel) into a water height map, which is a per-pixel water height in meters above sea level. The algorithm tries to find a physically reasonable water height map that best matches the input inundation map, where the physically reasonable requirement is defined as: (1) the water height surface must be smooth, i.e. we aim to find a water height map that does not change significantly between neighboring pixels; and, (2) the water height surface should not have a minimum or a maximum at the interior of flooded regions. This optimization problem is not differentiable, and thus cannot be easily solved directly. Instead, the following heuristic can be shown to produce an optimal solution to the above optimization problem. The algorithm identifies the boundaries of the inundated areas of the input inundation map. The water height at these boundaries is extracted from the corresponding DEM. In between these boundaries, the algorithm uses the Laplace differential equation to interpolate the water heights. The water height map is defined as a low-resolution image, where every pixel is set to be of 32x32 DEM pixels. This assures that the output map is smooth and does not contain high frequency changes, while also reducing the computational complexity of the process. In addition, outlier DEM pixels, which are pixels that cause high Laplace tension, are removed to assure that the overall function is smooth.

Gauge Stage To Flood Depth Algorithm

When inferring inundation depth for a real-time gauge water stage forecast, we do not have access to the full current flood extent (as the extent-to-height algorithm above assumes). To be able to provide depth in this more challenging setting, we first perform some precomputation. We first apply the Thresholding model described above to all past events in our training data, producing an inundation extent map for each gauge stage measurement. We then apply the flood extent to water height algorithm described above to produce a water height map for each past event.

In real-time, when we receive an input water stage at the target gauge, the model simply performs per-pixel piecewise-linear interpolation between the water height maps of the training dataset to generate a new water height map corresponding to this input stage. The resulting water height map and the DEM are then used to generate: (1) an updated inundation extent map, by assigning a dry state to a pixel if its water height is lower than the DEM and a wet state otherwise; and, (2) an inundation depth map, as the difference between the water height and the DEM height for wet pixels. The use of the Thresholding model (as opposed to directly using satellite-based imagery of actual historical flood extents) as the input flood extent maps ensures that higher gauge stages always yield larger inundation extents, and thus removes unnecessary noise. When the model infers an inundation depth map for a gauge’s water stage higher than all the events in the training set, it extrapolates the water height map by adding the gauge level difference to every pixel in the highest water height map computed from observed gauge stages, and uses the extrapolated water height and DEM to compute the water depth map.

Figure 5 illustrates the schema of the Manifold model.

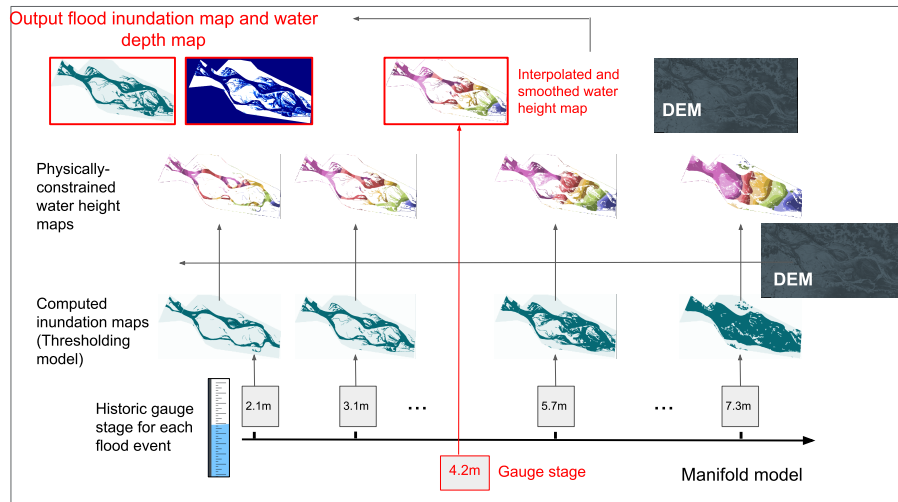


Fig. 5. Schema of the Manifold model.

Training and Validation The inundation models are trained and validated based on historical flood events, where flood inundation extent maps from satellite data, along with the corresponding gauge water stage measurements, are available. Similar to the stage forecast models, a 1-year leave out cross validation scheme is used for training and validation. For operational use, the models are retrained with all historical data. It should be noted that, contrary to flood inundation extent, it is much more difficult to obtain ground truth data for flood inundation depth at scale. The Manifold model is, therefore, trained and validated only on the inundation extent. However, since the Manifold model is constrained to only produce physically reasonable water height maps, accurate inundation extent metrics on the test dataset imply reasonably reliable inundation depth results (with further validation ongoing).

4 Resource Requirements

4.1 Data sets

The system utilizes the following data sets.

Precipitation Measurements We currently use IMERG satellite data, though other sources can be used (as long as the models are re-trained for the new dataset).

Precipitation Forecasts We currently use Global Forecast System (GFS) precipitation forecasts, though other sources can be used (as long as the models are re-trained for the new dataset).

Stream Gauges Stream gauges are measurement devices that measure the water level (and sometimes estimate discharge) of the river. We use these as labels for the hydrologic model, and inputs for both the hydrologic model and the inundation model. We use both open stream gauge datasets such as GRDC, CAMELS, and Caravan - as well as proprietary datasets from the governmental hydrometeorology agencies we work with.

Elevation Maps Accurate elevation maps are crucial for the inundation model, especially when producing depth estimates. The publicly available global DEMs (e.g., NASA SRTM and MERIT) unfortunately lack the required spatial accuracy and resolution for detailed flood inundation simulation. Furthermore, they are based on data from over a decade ago, thus failing to capture the frequent topography changes caused by past floods. Consequently, we currently construct higher-resolution, up-to-date DEMs for each AOI from high resolution satellite optical imagery data in a process that is based on stereographic imaging. To keep the DEM up to date, the model is retrained annually based on fresh imagery in locations where flooding causes frequent topography changes. We currently use proprietary Google-generated elevation maps, yet one can use any reliable hydrologically-conditioned elevation maps of sufficient resolution, which depending on the basin could be anywhere between 0.5 to 30 meters.

Historical Flood Extent Maps We use the synthetic aperture radar ground range detected (SAR GRD) data from the Sentinel-1 satellite constellation to determine flood inundation maps at known timepoints and locations [24]. At any AOI, a SAR image is available once every several days, from which an inundation map was inferred using a binary classifier. Every pixel within a SAR image is classified as wet/dry via a Gaussian-mixture based classification algorithm. In order to calibrate and evaluate the classification algorithm, we have collected a dataset of Sentinel-2 multispectral images of flood events that coincide with the SAR image dates and locations. Reference Sentinel-2 flood maps were created by calculating per-pixel Normalized Difference Water Index ($NDWI = (B3 - B8) / (B3 + B8)$), where B3 and B8 are green and near infrared bands, respectively) and applied a threshold of 0.

4.2 Computational Resources

Training and using these models requires computational resources as well. However, significant effort has gone into minimizing the computational requirements.

Training the hydrologic models takes seconds for the linear model and approximately 3 GPU hours for the LSTM (per location). Inference time is negligible for both.

Training the inundation models takes between 10 CPU minutes to 10 CPU hours depending on the AOI for both the lightweight model and the manifold model. For comparison, calibrating hydraulic models takes about 100-1200 CPU years for the same areas of interest [9].

5 Field Evaluation

At the time of writing (early 2022), these models have been incorporated in operational systems for several years, currently covering about 360 million people across India and Bangladesh. So far, the models' performance in real-time operational systems is consistent with their cross-validated performance on historical data.

5.1 Hydrologic Models

Arguably the most popular metric for evaluating the accuracy of hydrologic models is the Nash-Sutcliffe Efficiency coefficient. This metric is the R^2 score of the predicted discharge, where the R^2 score (or coefficient of determination) of a prediction is defined as $1 - \frac{\sum(\hat{x}_i - x_i)^2}{var(x)}$ (with x_i representing the true value of the i 'th example, and \hat{x}_i representing the prediction for the i 'th example). This can be thought of as describing the percentage of target's variance explained. An R^2 score of 1 means a perfect prediction (and no variance remaining), while a score of 0 means the prediction is no better than guessing the average. This trait makes this score easy to understand, regardless of the scale and behavior of the target variable. Since most of our models predict water level rather than discharge, we use the closest parallel - an R^2 score over the water level predictions. Our hydrologic models have achieved an R^2 score of about 0.98 in our operational systems.

A metric which is less common in the professional literature but is perhaps easier for non-hydrologists to parse is simply the average error of our predictions, in centimeters. Our hydrologic models have so far seen an average error of 8.5 centimeters, across all our operational systems.

5.2 Inundation Models

All of our models are evaluated at a 16x16 meter resolution. The precision of an inundation forecast is defined as the number of pixels that were both forecasted as being inundated and were indeed inundated divided by the number of pixels that were forecasted as being inundated. The recall of an inundation forecast is defined as the number of pixels that were both forecasted as being inundated and were indeed inundated over the number of pixels that were actually inundated. The f1 score of a forecast is the geometric mean of its precision and its recall.

Our inundation models have been run and evaluated at a resolution of 16x16 meters per pixel, and achieve 0.69 per-pixel f1 score on average across our operational systems.

5.3 Impact Evaluation

In addition to evaluating our systems' accuracy, we also aim to evaluate their effectiveness in driving protective action. In 2021, a randomized controlled trial

run by the Yale Economic Growth Center found that our collaboration with local NGOs to generate and distribute flood warnings in India increased the portion of people who received warnings prior to flooding events by 3x, and increased the portion of people who took protective action ahead of a flooding event by 4.4x.

6 Lessons Learned

In the five years since we began this project, we have learned many lessons along the way. Here are a few highlights.

Focusing on the model is often not enough to drive real-world change. The real world is complicated, and the effectiveness of systems depends not only on the quality of your model. You have to ensure you have good (real-time) data collection systems, build a good product around your model outputs, and build partnerships with the key stakeholders that would use your systems. Especially in underserved communities, many components of the service pipeline may be flawed - and so it's critical to monitor and ensure that the full process from model to on-the-ground value works properly.

Relatedly, missing data and data errors drive the majority of operational inaccuracy. Academic benchmarks and analyses will focus exclusively on model quality, yet in the vast majority of operational systems errors in model inputs can completely dominate over model mistakes. As a result, investing in error detection, error correction, uncertainty estimation, out-of-distribution detection, review and oversight, and other similar tools is incredibly important.

Interaction with floods is incredibly diverse and requires different communication in different locations. People have different needs and expectations from flood forecasting systems. For example, in areas that rarely experience flooding, a flood with a depth of 15 centimeters can be a critical event - and people will expect a severe warning for it. However, in areas that experience frequent flooding, shallow flood waters are considered a non-event and people may be frustrated or surprised by even an informational update for such events. In extreme cases, like in some of the most flood-affected regions in Bihar, India, people ask to be warned only if and when flood waters are expected to reach waist height. Another example of such diversity is how people communicate with others about floods. In some regions, such as in much of South Asia, communities will share with their networks any information they have about upcoming floods. In others, such as parts of South America, there is stigma associated with being affected by floods, and so people won't discuss it - including critical safety information.

References

1. World Bank Global Assessment Report on Costs and Benefits of Early Warning Systems. https://www.preventionweb.net/english/hyogo/gar/2011/en/bgdocs/Rogers_&_Tsirkunov_2011.pdf (2011), [Online; accessed 30-09-2018]

2. World Health Organization (WHO) - Global Report on Drowning. http://www.who.int/violence_injury_prevention/global_report_drowning/Final_report_full_web.pdf (2014), [Online; accessed 30-09-2018]
3. The human cost of disasters: an overview of the last 20 years (2000-2019). <https://www.undrr.org/publication/human-cost-disasters-overview-last-20-years-2000-2019> (2015), [Online; accessed 25-03-2022]
4. United Nations Office for Disaster Risk Reduction (UNISDR) - The Human Cost of Weather Related Disasters. https://www.unisdr.org/2015/docs/climatechange/COP21_WeatherDisastersReport_2015_FINAL.pdf (2015), [Online; accessed 30-09-2018]
5. The Centre for Research on the Epidemiology of Disasters (CRED) - Natural Disasters 2017. https://cred.be/sites/default/files/adsr_2017.pdf (2017), [Online; accessed 30-09-2018]
6. Adapt Now: A global call for leadership on climate resilience. https://files.wri.org/s3fs-public/uploads/GlobalCommission_Report_FINAL.pdf (2019), [Online; accessed 25-04-2022]
7. F-Score. <https://en.wikipedia.org/wiki/F-score> (2022), [Online; accessed 10-08-2022]
8. de Almeida, G.A.M., Bates, P., Freer, J.E., Souvignet, M.: Improving the stability of a simple formulation of the shallow water equations for 2-d flood modeling. *Water Resources Research* **48**(5) (2012). <https://doi.org/https://doi.org/10.1029/2011WR011570>, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011WR011570>
9. Ben-Haim, Z., Anisimov, V., Yonas, A., Gulshan, V., Shafi, Y., Hoyer, S., Nevo, S.: Inundation modeling in data scarce regions. *Neural Information Processing Systems (NeurIPS), Artificial Intelligence for Humanitarian Assistance and Disaster Response workshop* (2019)
10. Davenport, F.V., Burke, M., Diffenbaugh, N.S.: Contribution of historical precipitation change to us flood damages. *Proceedings of the National Academy of Sciences* **118**(4), e2017524118 (2021). <https://doi.org/10.1073/pnas.2017524118>, <https://www.pnas.org/doi/abs/10.1073/pnas.2017524118>
11. Doerr, S.H., Santín, C.: Global trends in wildfire and its impacts: perceptions versus realities in a changing world. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**(1696), 20150345 (2016)
12. Doocy, S., Daniels, A., Murray, S., Kirsch, T.D.: The human impact of floods: a historical review of events 1980-2009 and systematic literature review. *PLoS currents* **5** (2013)
13. Emerton, R.E., Stephens, E.M., Pappenberger, F., Pagano, T.C., Weerts, A.H., Wood, A.W., Salamon, P., Brown, J.D., Hjerdt, N., Donnelly, C., et al.: Continental and global scale flood forecasting systems. *Wiley Interdisciplinary Reviews: Water* **3**(3), 391–418 (2016)
14. Gauch, M., Kratzert, F., Klotz, D., Nearing, G., Lin, J., Hochreiter, S.: Rainfall-runoff prediction at multiple timescales with a single long short-term memory network. *Hydrology and Earth System Sciences* **25**(4), 2045–2062 (2021)
15. Hallegatte, S., Green, C., Nicholls, R.J., Corfee-Morlot, J.: Future flood losses in major coastal cities. *Nature climate change* **3**(9), 802 (2013)
16. Jongman, B., Ward, P.J., Aerts, J.C.: Global exposure to river and coastal flooding: Long term trends and changes. *Global Environmental Change* **22**(4), 823–835 (2012)

17. Jonkman, S.N.: Global perspectives on loss of human life caused by floods. *Natural hazards* **34**(2), 151–175 (2005)
18. Jonkman, S.: Loss of life caused by floods: an overview of mortality statistics for worldwide floods. DC1-233-6 (2003)
19. Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S., Nearing, G.: Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets. *Hydrology and Earth System Sciences* **23**(12), 5089–5110 (2019). <https://doi.org/10.5194/hess-23-5089-2019>, <https://hess.copernicus.org/articles/23/5089/2019/>
20. Kratzert, F., Klotz, D., Herrnegger, M., Sampson, A.K., Hochreiter, S., Nearing, G.S.: Toward improved predictions in ungauged basins: Exploiting the power of machine learning. *Water Resources Research* **55**(12), 11344–11354 (2019)
21. Loster, T.: Flood trends and global change. In: *Proceedings IIASA Conf on Global Change and Catastrophe Management: Flood Risks in Europe* (1999)
22. OECD: *Financial Management of Flood Risk* (2016). <https://doi.org/https://doi.org/https://doi.org/10.1787/9789264257689-en>, <https://www.oecd-ilibrary.org/content/publication/9789264257689-en>
23. Pilon, P.J., et al.: Guidelines for reducing flood losses. In: *Guidelines for reducing flood losses*. Naciones Unidas (1998)
24. Schumann, G.J.P., Moller, D.K.: Microwave remote sensing of flood inundation. *Physics and Chemistry of the Earth, Parts A/B/C* **83**, 84–95 (2015)
25. Subbiah, A., Bildan, L., Narasimhan, R.: Background paper on assessment of the economics of early warning systems for disaster risk reduction. World Bank Group for Disaster Reduction and Recovery (2008)
26. Zhongming, Z., Wei, L., et al.: The global climate 2001-2010: a decade of climate extremes-summary report (2013)