Regional Index Insurance using Satellite-based Fractional Flooded Area

Beth Tellman[1], Upmanu Lall[2], Saiful Islam[3], Md. Ariffuzaman Bhuyan[4]

1. University of Arizona, School of Geography, Development, and Environment
2. Columbia University, Department of Earth and Environmental Engineering
3. Bangladesh University of Engineering and Technology
4. Bangladesh Flood Forecasting and Warning Center

Corresponding author: Beth Tellman (btellman@arizona.edu)

Key Points:

1. Estimating return periods for fractional inundated area with the oft-used GEV (Generalized Extreme Value) distribution is inappropriate

2. Bounded distributions (e.g. Beta) reduce uncertainty estimates for probability of exceedance from short duration inundation time series

3. Example design and price of parametric flood insurance with fractional inundated area in Bangladesh and Argentina accounting for uncertainty

Plain language summary:

Index insurance, catastrophe bonds, and other types of risk transfer instruments could play an important role in adapting to floods and ensuring sustainable development in a world of increasing flood risk. In this article, we examine how satellite time series of inundation can be used to develop an emerging type of flood insurance, known as parametric or index-based insurance. Unlike traditional indemnity insurance, which relies on adjusters to estimate loss for individual damage, index insurance uses data ex-ante to determine payout contracts when pre-specified thresholds are crossed. Inundation extent from satellite remote sensing may provide a more direct measure of flood risk than data from models or stream gauges. However, typical methods used to estimate return periods for floods from models and gauges are not appropriate for fractional inundated area measurements from satellites. Here we provide a more appropriate method to estimate return periods and quantify uncertainty to price an insurance product leveraging the relatively short satellite record. Example applications for Bangladesh and Rio Salado, Argentina are provided. We show why estimating and pricing uncertainty ultimately influences insurance contract pricing and can help governments select insurance policies that align with their flood adaptation strategy.

Abstract:
Emerging parametric insurance products targeted at regional governments consider an index of flooding as the instrument for payoff and rate setting. Inundation extent from satellite remote sensing may provide a more direct measure of flood risk in this context than hydraulic modeling of flow and inundation. Here, we examine satellite-based fractional inundated area as a proxy for flood impact that can be used for index insurance payment at a regional scale. Typical methods for estimating return periods from unbounded distributions such as the GEV (generalized extreme value distribution) are not appropriate for fractional flooded area, which is bounded by 0 and 1. Here we examine alternative bounded distributions (2 parameter and a 4 parameter Beta) to estimate return periods and quantify uncertainty using a bootstrap sampling procedure for the short duration satellite record of fractional flooded area. We consider two examples with distinct flood dynamics i) a country (Bangladesh) where a flood can cover the majority of the land surface, and ii) a river basin (the Rio Salado basin in Argentina) where the worst flood covered only a modest fraction of the watershed. We explore how a parametric insurance policy based on fractional flooded area could be priced based on a typical approach used in the industry, that accounts for uncertainty for small sample estimation. Our exploratory approach to model selection illustrates how estimating the uncertainty price influences insurance contract pricing and is important to consider the choice of distribution beyond just the traditional measures of goodness of fit.

Index Terms: inundation, insurance, remote sensing, probability, risk transfer

1 Introduction

Flood frequency and magnitude estimates that underpin spatial zoning and risk mitigation strategies rely primarily on flood exceedance probabilities from discharge measurements estimated from gauges, models, or space-borne observations. Instead of indirectly estimating asset exposure through discharge or modeled inundation, satellite observations of inundated area can provide a more direct measurement of flood exposure. Given the growing length of the

satellite record, time series of inundated area could be used for exceedence probability estimation. Here we present the first exploration of whether inundation data could be used directly for parametric insurance products that seek to use an index of large area inundation as a mechanism for insurance and payment. The primary contribution of this paper is to open a new avenue for flood risk analysis and its mitigation, that can leverage the growing satellite record and complement traditional discharge and flood model-based approaches. We explore applications at the basin and country scale. Our application of interest is the financial securitization of catastrophic flood risk through a parametric instrument (an index insurance contract is used as the example, but others such as a catastrophe bond could also be considered) issued by a regional government seeking rapid funds in the event of disaster relief. Such a catastrophe may be indexed to the fraction of the area flooded in an event. The inundated area is bounded by the drainage area of a river basin, hydrologically, and by the area of the country, politically, so a bounded distribution is necessary whether one considers area inundated or fractional area inundated. This is a different setting than the one traditionally considered for flood risk where discharge is the target variable, and a fat right tail is considered appropriate for the extreme value distribution.

Flood frequency and risk analyses were developed in an engineering context where discharge was seen as a mass conserving and stationary process (as opposed to stage, which may be affected over time by sediment or erosion). Most literature regarding flood exceedance probabilities focuses on discharge measurements, usually from one or more points on a river (e.g. a gauge), or estimated from a rainfall-runoff model (Singh 2017). Physics-based models estimate velocity and stage as primary state variables (to calculate discharge) using a variety of land surface data, and parameterizations of key terms in the mass, momentum and energy conservation equations. Often, the parameters of these relationships need to be calibrated from at site data, which may be poorly constrained or even non-existent. Uncertainties associated with spatial inundation can consequently be large, especially in regions with limited topographic relief, or with complex river channel geometry, such as those in river deltas (Merwade et al. 2008, Teng et al. 2017, Bates 2022).

Stream gauge records are limited for estimating the impact of extreme flood events for several reasons. First, most gauges cannot measure extreme discharge and it is thus extrapolated using stage-discharge relationships. Second, declines in funding and civil conflict have led to a precipitous decline in global stream gauge data (Hannah et al. 2011). Third, many of the flood-prone areas in large river basins cross international boundaries where politics prevent sharing gauge data information needed to properly calibrate models (e.g. Mekong, Ganges Brahmaputra-Meghna, and Nile- see (Gleason and Hamdan 2017). Recent advances to estimate discharge via satellite are promising, but it remains difficult to estimate discharge of extreme flood events (Gleason and Durand 2020, Allen et al. 2020). Extreme events can influence sediment load and river conveyance, with implications for how bank full discharge volumes may change over time, influencing overbank flows and floods (Sofia and Nikolopoulos 2020). Changes in land use, demographics, and infrastructure, especially due to rapid urbanization of over 9,000 km^2 annually (Liu et al. 2020) also influence the degree to which discharge may induce different levels of flood damage. Human modifications of rivers (e.g. dam construction, paving floodplains, and channelizing rivers) mitigate, exacerbate, or shift flood waters (Chin 2006, Grill et al. 2019). Finally, discharge measurements for extreme flood events and physical estimates of them are themselves highly uncertain (Smith et al. 2018) given stage-discharge relationships that are calibrated to much lower flow regimes and for very different flow domains. Further, river

stage is only indirectly related to flood damage, which is more directly related to whether or not a location is inundated.

To relate discharge to flood damage, hydraulic models take discharge as an input to spatial estimates (often with considerable uncertainty) of extent and depth of inundation for discharges whose return periods have been estimated at a few index gauges. These are then used to generate population and property exposure estimates. Modeled inundation footprints are uncertain, in part because discharge measurements are the main input to models predicting inundation, and in part due to the changing geometry of the flow domain and potentially the changing density of flow as related to increasing sediment uptake, during extreme floods (Sofia and Nikolopoulos 2020). Flood damage estimation requires information on asset attributes, as well as on depth, flow and duration of inundation, and often has even higher uncertainty when predicted from the few attributes on which data is available (Merz et al. 2010, 2013).

Flood models also have difficulty predicting dam failure, levee breaches, failure in informal urban drainage systems, avulsions, and other extreme events, that are rare by definition (Swain et al. 2020). The accuracy of modeled global flood exposure is limited by inadequate data on flood defenses, topography, and discharge or flood footprints for calibration data (Ward et al. 2015). Flood models rely on computerized representations of watersheds that rarely include all human alterations to the earth's surface that impact flood dynamics- from fish canals in Cameroon (Shastry et al. 2020) to urbanization in Houston (Sebastian et al. 2019). Global hazard models differ in their structure and assumptions, leading to high disagreement of population and area exposure estimates (Ward et al. 2015, Trigg et al. 2016, Aerts et al. 2020). While incorporating high-resolution data improves flood model performance in the USA (Wing et al. 2017), these detailed elevation data are largely unavailable globally, and this is but one component of the uncertainty in model estimates.

Inundation is even more difficult to model than discharge for extreme floods because it requires modeling out of bank flow, and the data required to calibrate these models is sparse for extreme events. Remote sensing and social media have emerged to aid model calibration and to improve modeled floodplains (Bernhofen et al. 2018, Hultquist and Cervone 2020, Liang and Liu 2020, Scotti et al. 2020). Modeled inundation underpins existing insurance mechanisms, such as the FEMA 100-year floodplain maps to price the National Flood Insurance Program. However, these models are often based on estimation of 100-year discharge at ungauged locations, amplifying the associated uncertainty in inundated area. Yet these uncertainties are ignored in flood maps, leading to considerable basis risk, or when actual losses do not match predicted losses based on an index or data, for both the insurer and the insured.

These issues lead us to explore whether a more direct measure of inundation would be suitable for creating and verifying an index for flood insurance. For the past several decades, near-daily observation of changes in surface water and inundation are available through a variety of satellite sensors. From coarse resolution (25km) L-band radar capable of detecting water through vegetation on a daily basis since 1992 (Jensen and Mcdonald 2019), to daily high resolution (3-5m) optical observations capable of detecting water in dense urban areas (e.g., since 2017 from Planet), flood observation has become easier over time with an ever-increasing number, resolution, and type of sensor (Finer et al. 2018). Advances in parallel computing (e.g. Google Earth Engine) facilitate these advances. Several methods exist to estimate flood incidence and flood events from daily sensors at moderate resolution (MODIS, 250m) (Klein et al. 2015, Kuenzer et al. 2015, Policelli et al. 2016, Tellman et al. 2021a). Spatial flood frequency observation from the 20-year record MODIS satellite and nearly 40-year record Landsat satellite

can aid watershed agencies and governments in estimating relative areas of high risk (Hawker et al. 2020, Tellman et al. 2021b). These observations have been used, for example, to relocate refugee populations (Zajic 2019) in countries where flood models are unavailable at high resolution or have high degree of uncertainty (Trigg et al. 2016). Remote sensing of flood extent has inherent and known uncertainties, from vegetation, burned area, and clouds obfuscating the view of some pixels. However even accounting for these uncertainties, a recent study that compared satellite-based flood frequency have found that flood models largely overestimate flooding at low return periods (20 years or less), but were similar near the 50- or 100-year return period (Hawker et al. 2020). This study, however, was unable to test congruence or lack thereof between models and remote sensing data at higher return periods given uncertainty in flood models and remote sensing data. Leveraging the time series of remote sensing estimates of inundated area directly has so far not been explored for its potential insurance applications.

Parametric or index-based insurance is relatively new (Surminski and Oramas-Dorta 2014, Surminski et al. 2016). The key components of index insurance are the selection and verification of a payout index for the hazard of interest (in this case, floods). Currently, most flood-related parametric products rely on rainfall over a certain area and duration passing a specified threshold (e.g. CCRIF, the (Caribbean Catastrophe Risk Insurance Facility 2015) or are tied to Sea Surface Temperature (ENSO) proxies for flooding (e.g., in Peru, see (Khalil et al. 2007). Rainfall based insurance requires long station records, which are often available. The spatial variability of rainfall is significant (Arnaud et al. 2002, Smith et al. 2004) and hence the rainfall record is often not representative of the locations flooded and insured, leading to high basis risk. Spatial variability in rainfall may be limited in events like Hurricane Harvey, with persistent and spatially uniform rainfall over a large area. However, for compound events such as the 5 typhoons that occurred in 90 days leading to widespread and persistent flooding in Thailand in 2010, how to trigger an index based on rainfall is unclear. Does the trigger payout for every typhoon? And which locations for that same year? In a country like Bangladesh, for example, where 70% of the watershed area and thus the rainfall leading to floods, is outside of the country borders, an index based on local rainfall in Bangladesh is unlikely to work and must be paired with discharge (Haraguchi 2018). All existing measurements- rainfall, discharge, or model-based inundation can lead to high basis risk. In all cases, the instrument defining the index needs to be measured transparently, and by a neutral third party. Satellites provide the ability for remote and transparent inundation measurement as both an insurance pricing and payoff mechanism. In Bangladesh, the first inundated area flood index insurance pilot based on the MODIS satellite has shown promising results (Matheswaran et al. 2019, Amarnath 2020). Here, we consider fractional inundated area could be computed with respect to a spatial unit (country, watershed, district, city) and serve as the index and trigger for insurance. The relatively short satellite records for inundation, coupled with significant climate variability, may lead to high estimation uncertainty for insurance pricing, and a key issue we explore in this paper.

Index insurance applications ideally require out of sample estimates of inundation return periods that are unbiased and have low uncertainty. Methods that lead to high uncertainty would cause the insurance instrument to be more expensive, to cover the potential risk to the insurer. The primary question we explore in this paper is how to choose an appropriate probability distribution to estimate the return periods of satellite-derived fractional flooded area, while quantifying both bias and uncertainty.

We address these questions in two contexts (Bangladesh and Rio Salado, Argentina) with different flood generating mechanisms and terrain. Riverine flooding in Bangladesh mainly

occurs in wide areas on flat terrain, driven by characteristics of the monsoon and major rivers that are influent into a delta (Islam et al. 2010). By contrast, the Rio Salado basin in Argentina, , has frontal, convective, and episodic precipitation occurring over areas with more topographic relief (Latrubesse and Brea 2009), and hence the potential for full basin area flooding is limited.

In this initial effort, recognizing the resolution and coverage of historical satellite inundation products, we consider a country wide insurance product (e.g. Bangladesh) or a river basin level insurance product (e.g., a basin in Argentina), that would be purchased by a government or a bank to securitize the risk at these larger scales. Such a product, when triggered, would enable rapid release of funds for relief and reconstruction if a catastrophic flooding event were to occur. In this initial work we do not consider individuals as insured parties, recognizing that a macro level product may manifest significant basis risk for those sub-regional users.

1.2 The role of uncertainty when pricing flood index insurance

Index insurance is designed to provide financial support for infrequent or catastrophic loss. Unlike traditional indemity insurance, which relies on adjusters to estimate loss for each invidual damage, index insurance uses data (e.g weather data or increasingly, based on satellites (Enenkel et al. 2019, Benami et al. 2021)) ex-ante to determine payout contracts when pre-specified thresholds are crossed. To ensure a high correlation between the index and loss outcomes, accurate probability distributions are essential to identify thresholds for payouts (Clement et al. 2018). Poor correspondence between the index and insured losses is known as basis risk. Basis risk can lead to policy holders not receiving payouts when a catastrophic event occurs, resulting in low demand of purchasing the insurance product. The insurer may also be concerned about too frequent payouts, and may thus increase the premium to cover that risk. If the insured is consistently paid more often than designed, leading to losses for the insurer, potential discontinuance of the product may result. There are multiple sources and causes of basis risk, including inadequate length of data, poor design of an index that was not validated by loss and damage data (Osgood et al. 2018), inadequate index measurement of spatial variance of hazard (Norton et al. 2012), and others (see a review in Benami et al. 2021). Here we focus on the basis risk due to uncertainty associated with computing the probability that an index will lead to payouts, and how this uncertainty affects the price of the contract (Benami et al. 2021).

We illustrate the structure for the index insurance instrument in Figure 1. Most index insurance contracts consider an initial payout threshold to trigger a payout, and a payout limit (exit), which is the maximum possible payout. For example, the sovereign risk-pooling insurance contract for Caribbean countries (CCRIF) typically considers an initial threshold corresponding to a 15 year return period (e.g. red line), and an exit threshold that may correspond to a 75 or a 200 year return period (e.g orange line) (Caribbean Catastrophe Risk Insurance Facility 2015). Some contracts may set continuously increasing payments up to the exit or maximum payout, while others may payout in a stepwise fasion, only when exceeding specific thresholds (e.g. corresponding to the 50 and 100 year return period events). The threshold probabilities and the payout structure translate into an expected value of the contract that is central to determining the premium price. Transaction costs, profits, and a risk premium to cover the uncertainty of return period estimates is added on to the expected value or "fair premium".

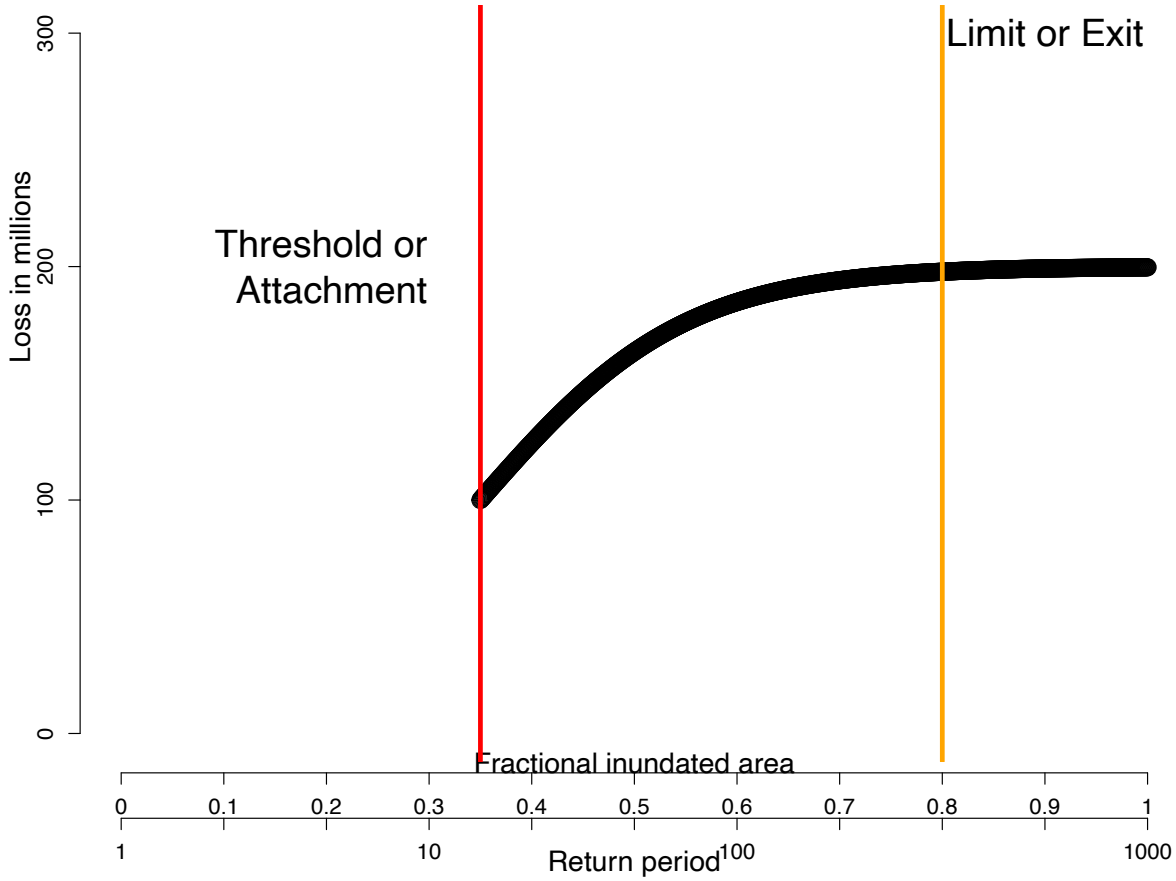

Figure 1. Potential parametric insurance loss structure of the initial payout (in orange), the payout limit (in red), and the loss curve for fractional inundated area, the associated flood return period, and expected damages (in black).

Many types of index insurance designs are possible, ranging from a single threshold and payout (e.g. the red line) to multiple interval payouts (red and orange lines, with other possible payout intervals along the black line in between). In all cases, one needs an estimate of the average exceedance probability of the threshold at which payoff occurs, and of the uncertainty associated with the estimate. Note that estimating the uncertainty for small samples (such as the relatively short satellite record) is addressed in section 2.3.

Consider a simple index insurance product that is designed as follows. A single threshold for payout corresponds to the fractional area inundated with annual exceedance probability p_{exc} (red line). It is essential to estimate the variation or uncertainty in the p_{exc} associated with the threshold or trigger and include it in the product price. This product is offered on an annual basis, and p_{exc} could potentially be re-evaluated annually based on either additional information available or using a predictive proxy index (e.g. (Khalil et al. 2007)). The simplest design of the product is that if the index threshold corresponding to p_{exc} is exceeded, then the insured party gets a fixed payout P. For simplicity, in this example let us consider that the terms are that a single payout per term of an insurance contract is considered. A common way to set a premium

is outlined in equation 1, where the premium is determined by the risk load, the profit margin for the insurer, and the transaction costs.

$$R_{x*} = P * p_{exc} + k * P * sp_{exc} + f1 + f2 * P \quad (1)$$

Here, R_{x*} is the premium for a fractional inundation threshold of x* (e.g., 0.3), p_{exc} is the corresponding estimate of the probability of exceedance, sp_{exc} is a measure of the uncertainty of the estimated p_{exc} (e.g., standard deviation of p_{exc} or the spread between the 95th and 5th percentiles of the p_{exc} associated with the trigger threshold based on bootstrapped samples, $|p_{exc,95\%} - p_{exc,5\%}|$), k (e.g., 5) is a risk pricing factor that is set by the company as a policy parameter reflecting their degree of risk aversion, f1 (e.g., \$100k) is a transaction cost, and f2 is a profit margin (e.g., 0.06) that is proportional to the payout amount, P. The key parameters of interest are then the estimates of the p_{exc} and the sp_{exc} corresponding to a specified threshold. The first term ($P * p_{exc}$) is the fair premium or the amount an insurer needs to charge the client to cover the payment, while the second term ($k* P * sp_{exc}$)is the risk premium, or the amount the insurer must charge to cover the uncertainty of the probability estimation. These are discussed next.

2 Materials and Methods

2.1 Choosing a distribution to model annual maximum fractional flood area

From the extensive experience in fitting flood frequency curves, it is known (Vogel 1986, Razali and Wah 2011) that for the typical sample sizes available for estimating the return periods of floods from streamflow discharge data, statistical tests such as the Kolmogorov-Smirnov test, the Chi-Square test and Filliben's correlation have low power in discriminating between multiple candidate distributions. Consequently, choices such as the Log-Normal or the Log Pearson III were adopted by committee in the USA (USGS 1982) and eventually, the generalized extreme value distribution (GEV). The GEV is an asymptotic distribution (i.e., as the sample size approaches ∞) for the block maxima of variables that may be drawn randomly from different distributions; a property which led to its emergence as a commonly used distribution (Chowdhury et al. 1991, Martins and Stedinger 2000). Douglas and Vogel (2006) show that the GEV, Log Normal, and Log Pearson 3 distribution can be a good fit for annual maximum floods in the USA and in the UK, while the Gumbel distribution is widely used in India. However, for most typical applications, the GEV is unbounded on the upper tail. When the variable of interest is the annual maximum of the fraction of area inundated in the country or basin, the application of the GEV is not appropriate, since this variable is bounded by 0 and 1. Botero and Francés (2010) found GEV was unacceptable to estimate distributions with an upper bound, e.g. when a physical limit applies, such as a PMP (Probable maximum precipitation) or PMF (probable maximum flood). Botero and France (2012) provide alternatives to GEV estimation, including the 4 parameter log-normal and 4 parameter extreme value distribution. Others have explored an inverted Weibull as an alternative (Bardsley 2018), which has the issue of lower values estimates going negative and changing the shape of the distribution.

The GEV distribution is defined as:

$$f(x) = \frac{1}{\sigma} t(x)^{\xi+1} e^{-t(x)} \quad where\ t(x) = \left\{1 + \xi \frac{(x-\mu)}{\sigma}\right\}^{-1/\xi} \qquad (2)$$

Where σ is a scale parameter, μ is a threshold (x > μ), and ξ is a shape parameter >0 and the domain of $x \in (0, \infty)$.

For the annual maxima of fractional flooded area, the variable of interest in this paper, the Beta distribution is an attractive candidate since it is bounded, and can also take a variety of shapes using only 2 parameters. We explore the 2 parameter Beta distribution, bounded at 0 and 1. A generalization of the Beta distribution that considers arbitrary bounds, not restricted to 0 and 1, can also be considered (e.g. (Wang 2005). However, given the short remote sensing records that are available (~20 years), estimating 4 parameters (the two shape parameters, and the 2 bounds) may significantly increase the uncertainty of return period estimation.

The 2-parameter Beta distribution is defined as:

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} \qquad (3)$$

Where x is a random variable, $x \in (0,1)$., and α and β are shape parameters of the distribution, and B(.,.) is the Beta function.

The 4-parameter Beta distribution is defined as:

$$f(x) = \frac{(x-\min)^{\alpha-1}(max-x)^{\beta-1}}{B(\alpha,\beta)(max-\min)^{\alpha+\beta-1}} \qquad (4)$$

Where of $x \in (min, max)$.

The estimation of the p_{exc} and the sp_{exc} is discussed in the next section.

2.2 Methods for return period estimation

The average probability of exceedance given a threshold is estimated as

$$p_{exc} = 1 - \int_{-\infty}^{X_T} \hat{f}(x)dx \qquad (5)$$

where $\hat{f}(x)$ is the probability distribution estimated from the sample data and p_{exc} is an estimate of the probability of exceedance. Many academics and practitioners instead routinely compute the average (expected value) of the flood discharge corresponding to a specified return period. This is typically expressed as $E\left[\widehat{X_T}\right] = \bar{x} + K_T s$, where T is the return period (=1/p), $\widehat{X_T}$ is an estimate of the corresponding quantile X_T, p is the probability of exceedance, $\bar{x}$ is the mean of the sample, s is the standard deviation, K_T is a frequency factor for the probability distribution model selected corresponding to the return period, and E[.] is the expectation or the averaging operator.

However, estimating the exceedance probability of a threshold (or quantile) (equation 5) is not equivalent to the latter approach of estimating the threshold (or quantile) of a specific exceedance probability. This issue was discussed extensively in the hydrologic flood frequency literature (Beard 1960, 1997, Tai 1987, Stedinger 1997). The sampling distribtions f($\hat{p}$) and f($\widehat{X_T}$) are asymmetric leading to $p_{exc} \neq \int_{-\infty}^{\widehat{X_T}} \hat{f}(x)dx$ (the asymmetry increases for smaller samples and for higher return periods).

An appropriate way to estimate the parameters of the distribution using the data is to maximize the likelihood of the sample relative to the candidate values of $\boldsymbol{\theta}$:

$$\max_{\boldsymbol{\theta}} L(\boldsymbol{\theta}) = \max_{\boldsymbol{\theta}} \prod_{i=1}^{n} f(x_i|\boldsymbol{\theta}) \quad (6)$$

Since (x_1, x_2, …..x_n) is a random sample, the resulting estimate $\widehat{\boldsymbol{\theta}}$ also has a probability distribution f($\widehat{\boldsymbol{\theta}}|\boldsymbol{x}$), as do the estimates of interest, f($\hat{p}_{exc}(x^*)|\widehat{\boldsymbol{\theta}}, x^*$) and f($\hat{x}_T|\widehat{\boldsymbol{\theta}}, T$), where ^ denotes the sample estimate. Since these probability distributions are typically asymmetric, the mean of the estimated probability of exceedance for a specified x* will not match the corresponding probability of exceedance (equation 7) for the expected value of the flood magnitude for that probability of exceedance. The same applies to the uncertainty distributions.

$$\mathrm{E}[\hat{p}_{exc}(x^*)|\widehat{\boldsymbol{\theta}}, x^* = x_T] \neq p_{exc}(E[\hat{x}_T|\widehat{\boldsymbol{\theta}}, T = 1/\mathrm{p}_{exc}]) \quad (7)$$

Placed in context of our flood parametric insurance example, if we estimate that the average return period associated with a particular level of fractional flooded area of 0.4, is for example, 100 years, then the average fractional flooded estimated corresponding to a return period of 100 years from a finite sample will not be 0.4. Typically using $\widehat{X_T}$ leads to a biased estimate of p_{exc} that for small samples can significantly understate the risk associated with that threshold (Beard 1960, Stedinger 1983). Once a payoff threshold is defined, the average exceedance probability $E[\hat{p}_{exc}]$ associated with that threshold informs the average payoff. Therefore for parametric insurance applications, estimating the exeedance probability of a threshold (not the opposite direction- estimating the threshold tied to a specific return period or exceedance probability) is preferred for unbiased results. Noting that the payoff occurs corresponding to a specified threshold, and for pricing we need to know its average probability of exceedance, we recommend (and construct an insurance price model) estimating the exceedance probability by fitting a distribution to the annual maximum inundation area data, and calculating uncertainty of this estimate via bootstrapping.

We illustrate the difference in estimating exceedance probability for a threshold/quantile (what we recommend) versus estimating the threshold or quantile for a specified exceedance probability for the Beta distribution with shape parameters equal to 1.5 and 4 respectively in Figure 2. The theoretical quantiles corresponding to return periods of 10, 20, 50, 100, 200, 500 and 1000 years were computed first. Subsequently, we draw 1000 samples of size 20 each from the distribution, using maximum likelihood estimation for the shape parameters, followed by an estimation of the quantiles for each of the return periods, and of the return period for each corresponding theoretical quantile. The theoretical quantile-return period relationship is illustrated using the symbol "o" in Figure 2. The blue line end points mark the 10th to 90th percentile of the return periods corresponding to each theoretical quantile, while the + symbols mark the 25th and 75th percentiles from the 1000 simulations. The red lines and +

symbols provide the corresponding information for the quantiles estimated for each return period. The asymmetry in the uncertainty distributions is notable. The uncertainty in the return periods is quite dramatic for this sample size. Uncertainty reduction by using a sample size of 200 is shown in the second part of the figure.

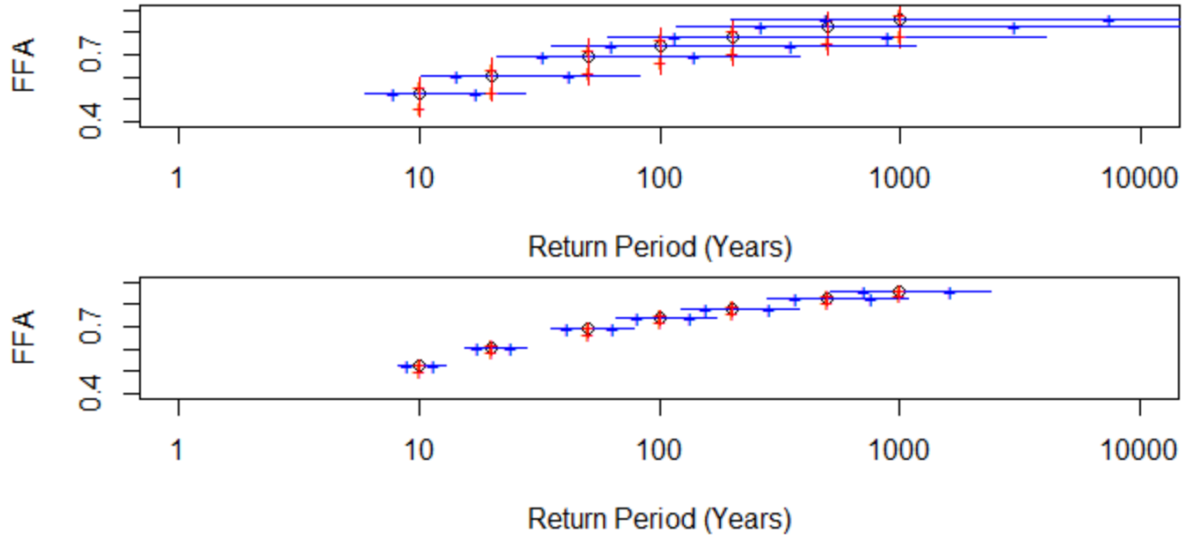

Figure 2. Sampling distributions of $\hat{p}_{exc}$ and $\hat{x}_T$ estimated using maximum likelihood for samples drawn from a Beta distribution with shape parameters 1.5 and 4, as an example for fitting a model to the fractional flooded area (FFA). Sample sizes are 20 (top) and 200 (bottom) respectively. Blue lines represent the uncertainty in the estimated return period and red lines in the estimated fractional flooded area corresponding to a quantile. The lines go from the 10th to the 90th percentile, The "true" quantile return period is shown using "o", and the 25th and 75th percentiles are marked with a "+".

The uncertainty of estimation of $\hat{p}_{exc}$ as represented by the difference between the 95th and the 5th percentile, is shown in Figure 3 as a function of the theoretical quantiles corresponding to each of the return periods, as a ratio of the average estimated $\hat{p}_{exc}$. Note that the relative uncertainty increases significantly as more extreme events are considered as would be expected. This shows that for the small samples typically available, the risk premium (term 2 in equation 1) could dominate the fair premium (term 1 in equation 1). When the ratio of the risk premium to the fair premium >1, then the risk premium dominates the price of the insurance instrument. The uncertainty ratio increases with fractional flood area (FFA), because large, infrequent, extreme floods are estimated with higher relative uncertainty.

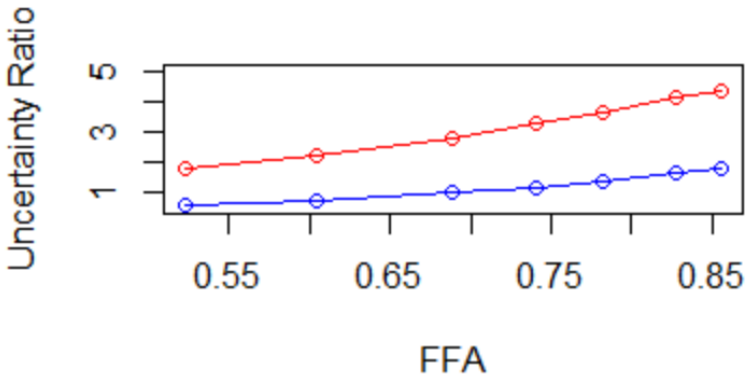

Figure 3. Ratio of estimated uncertainty to mean value of probability of exceedance from 10,000 bootstrap samples of size 20 (red) and 200 (blue) for the synthetic data drawn from a Beta distribution with shape parameters 1.5 and 4, with estimation by maximum likelihood, evaluated for fractional flooded area (FFA) corresponding to 10, 20, 50, 100, 200, 500 and 1000 year return periods. For small samples the risk premium (term 2 in equation 1) due to uncertainty will increasingly dominate the fair premium (term 1 in equation 1) as more extreme events are considered.

The parameter vector $\widehat{\boldsymbol{\theta}}$ for a candidate distribution is estimated by maximum likelihood. For the Beta distribution, the shape parameters are obtained through maximum likelihood estimation, using Rfast (v. 1.9.9 (Papadakis et al. 2020). For the 4 parameter beta distribution we use the ExtDist package in R (Wu, Haizhen et al. 2020) to estimate the minimum (min) and maximum (max) bound (instead of imposing it a priori as 0 and 1) (equation 3). Upper and lower bounds are solved using numerical maximum likelihood.

For the extreme value distribution we used the ExtRemes 2.0 package in R (Gilleland and Katz 2016) and the VGAM package (Yee 2010) to estimate the shape parameter of the GEV. In all cases we bootstrapped the time series with 10,000 random draws with replacement to estimate the mean and 95% confidence intervals for each statistic of interest, including $\hat{p}_{exc}$ and $\hat{x}_T$. We used the ppcc package (Pohlert 2020) to compare the Beta and GEV cumulative distribution to the empirical cumulative distributions using the probability plot correlation coefficient (ppcc) test. The estimates from bootstrapping the sample to estimate Beta and GEV parameters were used to compute the mean and confidence intervals for return periods to estimate the insurance premiums for each distribution (using equation 1).

2.3 Inundation Data for Bangladesh and Argentina

We used inundated area for two regions, a watershed and a country, summarized in Figure 4. The country scale data is from Bangladesh, a mega delta of three major rivers, which experiences consistent large area inundation. Riverine flooding in Bangladesh occurs mainly in wide areas on flat terrain, driven by characteristics of the monsoon and the major rivers that are influent into a delta (Islam et al. 2010). As the two largest flood events in Bangladesh history occurred before the daily satellite record from MODIS (in 1988 and 1998), we use annual

maximum inundated fractional area estimates of the country based on the Flood Forecasting and Warning Center (FFWC) flood model (MIKE11 from DHI (Danish Hydraulic Institute) from 1975-2019. The model uses 74 rainfall stations and 94 water level stations and a 300m spatial resolution DEM to estimate inundated area. The time series of inundated fractional area (Figure 4b) and maximum inundated area observed (the 1998 flood- Figure 4d) are displayed in Figure 4.

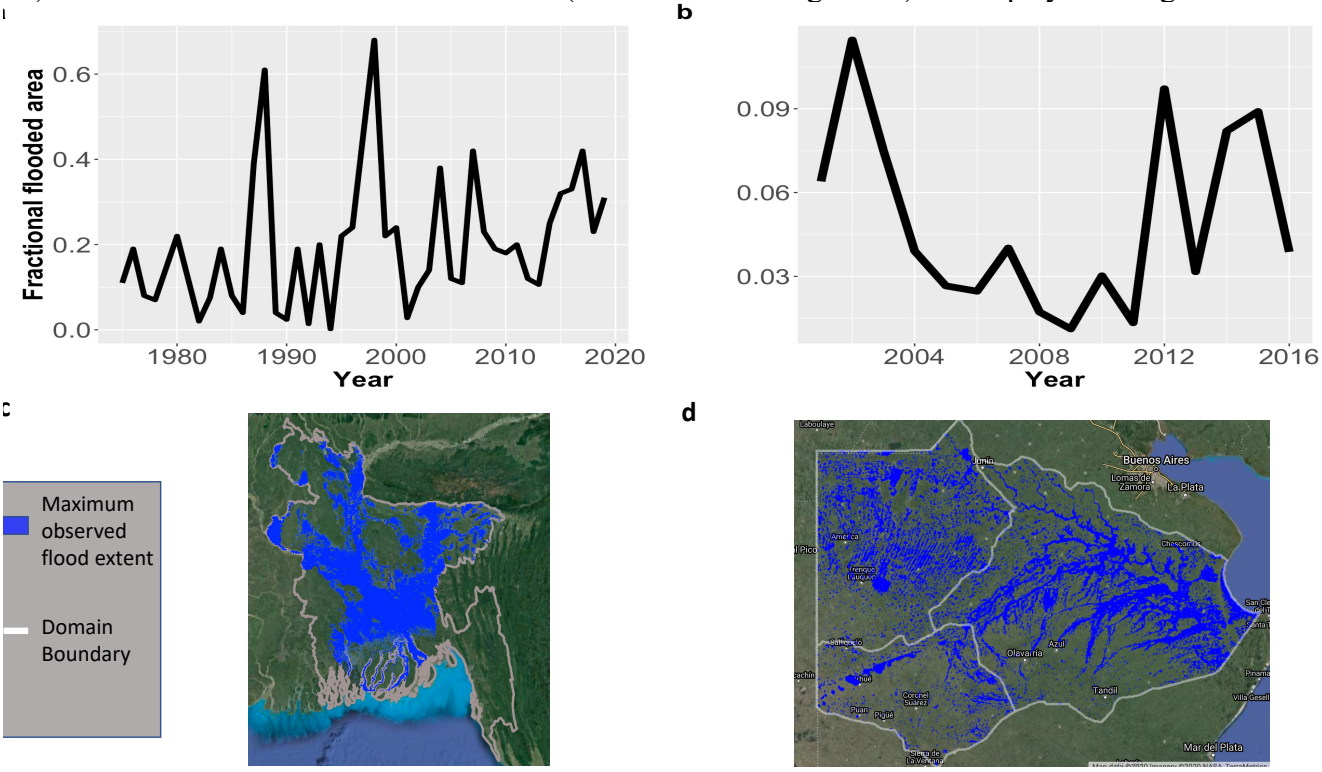

Figure 4. Time series of the fraction inundated and map of study area of maximum observed flood extent in each study region. a) Inundated area time series in Rio Salado, MODIS satellite derived, 2001-2016 (Tellman 2021c). b) Inundated area as percent of the country, Bangladesh, from the Flood Forecasting and Warning Center (FFWC) model, using the Mike 11 hydrodynamic model coupled to a Mike 11 RR rainfall runoff model 1975-2019. c) Rio Salado watershed, Argentina, with maximum observed flood extent over the time series, d) Bangladesh, maximum modeled flood extent for 1998, covering 68% of the country, from the FFWC.

Previous research suggests that severe damage to crops and infrastructure in Bangladesh occurs with 0.26-0.34 fraction of the country flooded, extensive damage occurs when 0.34-0.385 is inundated, and catastrophic damage with exceptional economic loss occurs at >0.385 fractional area inundated (Monirul Qader Mirza 2002). More recent damage data collected since 2010 (Table 1) indicates however the two largest and most extreme events were over a fractional area inundated of 0.6. A smaller flood that occurs at an inopportune time in the planting or harvesting season, or in an urban area, may cause more damage. Inundated area and damage relationships have likely changed over time, given the high rate of population growth and urbanization in floodplains in Bangladesh (Tellman et al. 2021a) and investments in flood control and embankments(Rahman and Salehin 2013). Ultimately a country would want to consider insurance for the inundated area that causes damage given its current land use and flood exposure conditions.

Table 1. Estimated damage in USD for extreme or catastrophic flood events reported in both Bangladesh and Rio Salado, adjusted for inflation and reported in 2010 USD.

Location	Year	Damage (USD) billions	Fractional inundated area	Source
Bangladesh	1974	4.5	0.36	Word Bank 2010
	1987	4.498	0.39	
	1988	4.884	0.61	
	1998	4.023	0.68	
	2004	2.941	0.38	
	2007	1.365	0.42	
	2017	.0671	0.42	(World Food Program 2017)
Rio Salado	2002	1.32	0.11	(Ibarlucía et al. 2017)
	2012	0.42	0.1	Ibarlucía et al. 2017)
	2015	0.072	0.09	(Rivas 2017)

The watershed-scale times series data is for the Rio Salado basin in Argentina. The Rio Salado Basin is 170,000 km^2 in the Buenos Aires Province in Argentina. Similar to Bangladesh, this low sloped basin (1/1000) experiences frequent wide-area inundation. The Rio Salado basin in Argentina, in contrast, has more frontal, convective, and episodic prediction occurring over areas with more topographic relief (Latrubesse and Brea 2009), and hence the potential for full basin area flooding is limited. Maximum annual flooded area was estimated from a remotely sensed inundation time series (Tellman et al. 2021b) using the MODIS satellite from 2000-2017. Only the data from 2001-2016 were used for analysis in this paper, because in the year 2000 only one (Terra) of the two daily MODIS sensors (Aqua and Terra) was operational. The 2001-2016 time series represents data of maximum annual inundation by applying an inundation algorithm to twice daily 250m resolution MODIS images to estimate pixel inundation. Notable damaging floods during this time occurred in 2002 ($700M USD), 2013 ($546M USD), and 2015 ($82M USD) (Ibarlucía et al. 2017) (see Table 1).

The inundation algorithm is a modified version of the Brakenridge and Anderson (2006) algorithm (see Tellman et al. (2021), summarized here). The algorithm uses the near-infrared (NIR) at 250-meters and short-wave infrared (SWIR) band (pan-sharpened from 500 to 250m) to monitor surface water. Pixels are flagged as water or non-water if their values are below thresholds across three bands: red (b1), SWIR (b7) and a ratio of NIR-red (b2b1ratio). Threshold values are based on Brakenridge and Anderson (2006) and determined from regression of discharge data on MODIS digital numbers (reflectance values scaled by 10,000 (E. Vermote and R. Wolfe 2015). Cloud and cloud shadow misclassifications are accounted for with a multi-day compositing technique that classifies a pixel as water if it is stable throughout a composite window. A 3-day composite is used where a pixel must be classified as water in 3 out of the 6 available images to remain classified. As clouds and cloud shadows move throughout the 3-day composite, overlapping misclassifications of cloud shadows do not occur frequently enough (e.g. 3 times in 6 images over 3 days) to provide a significantly large number of false-positive

observations. The daily water detections are combined into annual composites, representing a total maximum annual flood extent. The fractional flooded annual area of Rio Salado are presented in Figure 4A, with maximum observed historical inundation extent in Figure 4C.

4 Results

4.1 Comparative Analysis of model fitting for the two regions

Theoretically, the GEV is an inadmissible distribution to use for fractional inundated area, because it is not bounded at 1. We present it here to show the implications of blindly using this distribution given its popularity for flood risk analysis. We compare the 2 and 4 parameter Beta distribution to assess if the bound of the distribution is likely to be below 1 (which the 4 parameter Beta can estimate).

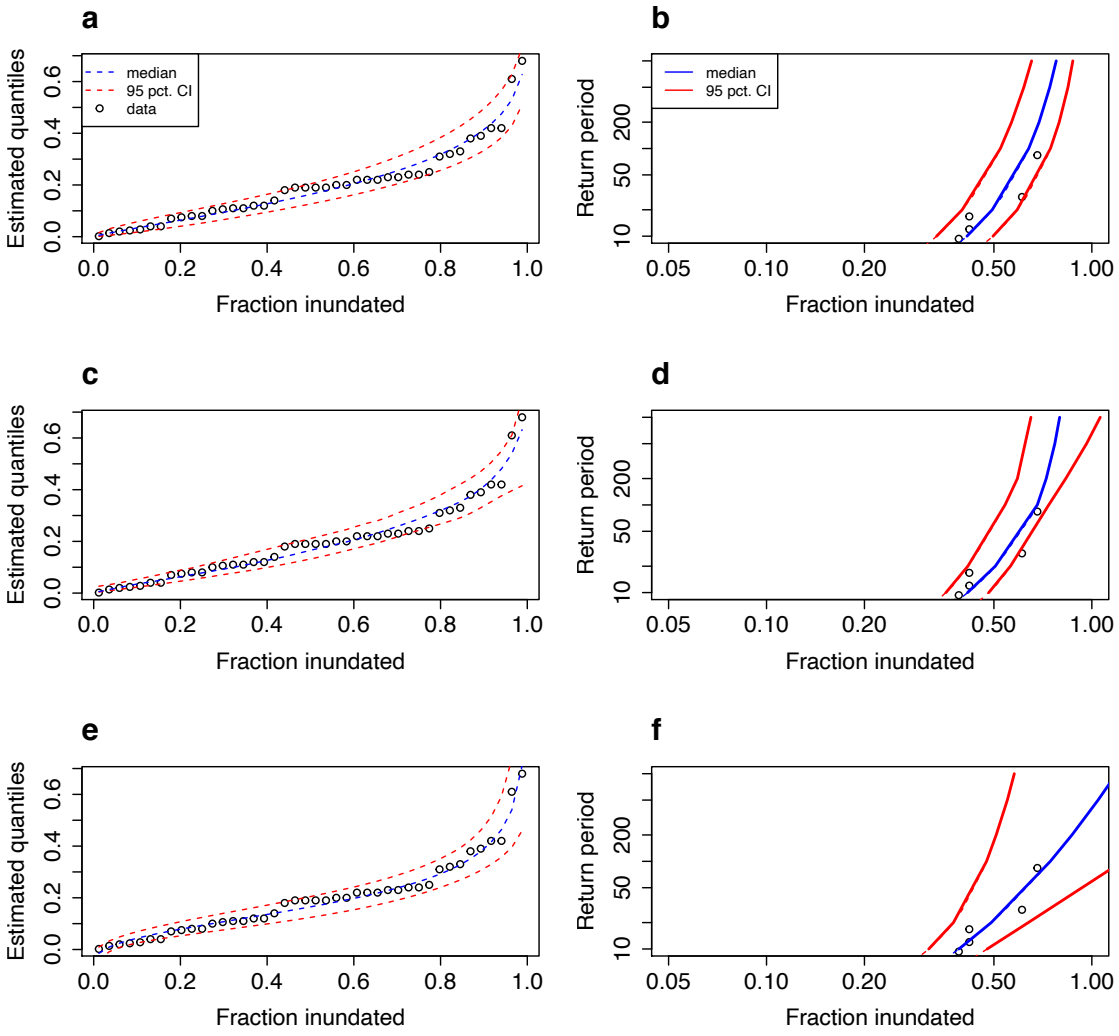

Figure 5. Beta and GEV comparisons for Bangladesh. a) 2 parameter Beta modeled quantiles; b) 2 parameter Beta distribution return period estimates and 95% confidence intervals of the quantiles; c) 4 parameter Beta modeled quantiles; d) 4 parameter Beta distribution return period estimates and 95% confidence intervals quantiles ; e) GEV modeled quantiles f) GEV

distribution return period estimates and 95% confidence interval quantiles. Black points are observed data.

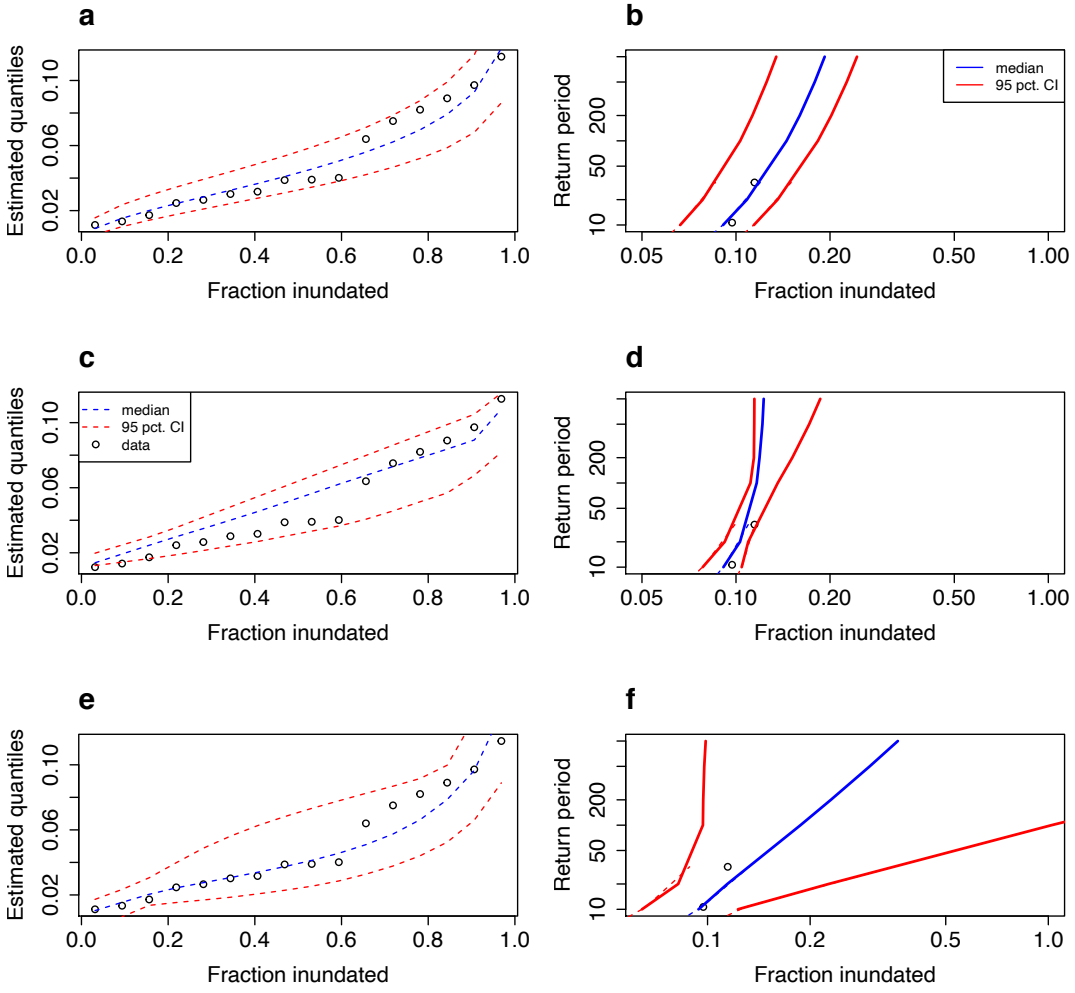

Figure 6. Beta and GEV comparisons for Argentina a) 2 parameter Beta modeled quantiles; b) 2 parameter Beta distribution return period estimates and 95% confidence intervals of the quantiles; c) 4 parameter Beta modeled quantiles; d) 4 parameter Beta distribution return period estimates and 95% confidence intervals of the quantiles; e) GEV modeled quantiles f) GEV distribution return period estimates and 95% confidence intervals of the quantiles. Black points are observed data.

The probability plot correlation coefficient tests (ppcc) do not reject the Beta distribution or the GEV for both the Argentina and Bangladesh data. The ppcc values for Argentina were 2 param Beta = 0.979, p=0.45; 4 param Beta= 0.977, p=;0.60 p=, GEV =0.946, p=0.32 and for Bangladesh 2 param Beta = 0.989, p=0.56; 4 param Beta= 0.984, p=0.80; GEV =0.99, p=0.32. This lack of ability to discriminate between plausible distributions is typical for the small samples available here. The corresponding quantile plots of modeled distributions also do not discriminate between both Beta distributions and the GEVwhen compared to observed data (Bangladesh Figure 5 a,c,e; Argentina, Figure 6 a,c,e).

However, the return period estimates reveal differences in uncertainty for each distribution. In Bangladesh, the GEV return period plot (Figure 5f) reveals that the 95% confidence interval exceeds 1 (a physically infeasible value, which does not even appear in the plot bounded at 1) at the 50-year return period. The 4 parameter Beta distribution (Figure 5d) also has greater uncertainty than the 2 parameter Beta distribution (Figure 5b) especially for return periods greater than 100 years.

In Argentina, the 4 parameter Beta distribution has the tightest uncertainty bounds (Figure 6d), compared to both the 2 parameter Beta (Figure 6v) and the GEV (Figure 6f). While the GEV does not exceed physically infeasible values as it did for the Bangladesh example, the uncertainty out of sample is very large.

PDFs of the upper bound of the 4 parameter Beta distribution reveals it approximates 1 in Bangladesh (Figure 7c). However, in Argentina, the upper bound is likely much lower than 1 (estimates cluster near 0.2) (Figure 7a), suggesting there may be a physical upper limit for this basin much lower than 1. Thus, the 4 parameter Beta distribution indicates lower uncertainty in the Argentina case, despite having to estimate 2 additional parameters. However, in the bootstrap samples there appear to be 17 out of 1000 simulations that select the upper bound of 1 (Figure 7b). There are two peaks in the distribution of the upper bound estimate for Argentina because we used a Bayesian procedure that selects and upper bound of 1 (what would result if a 2 parameter distribution were used) part of the time and a Beta distribution with a lower bound part of the time.

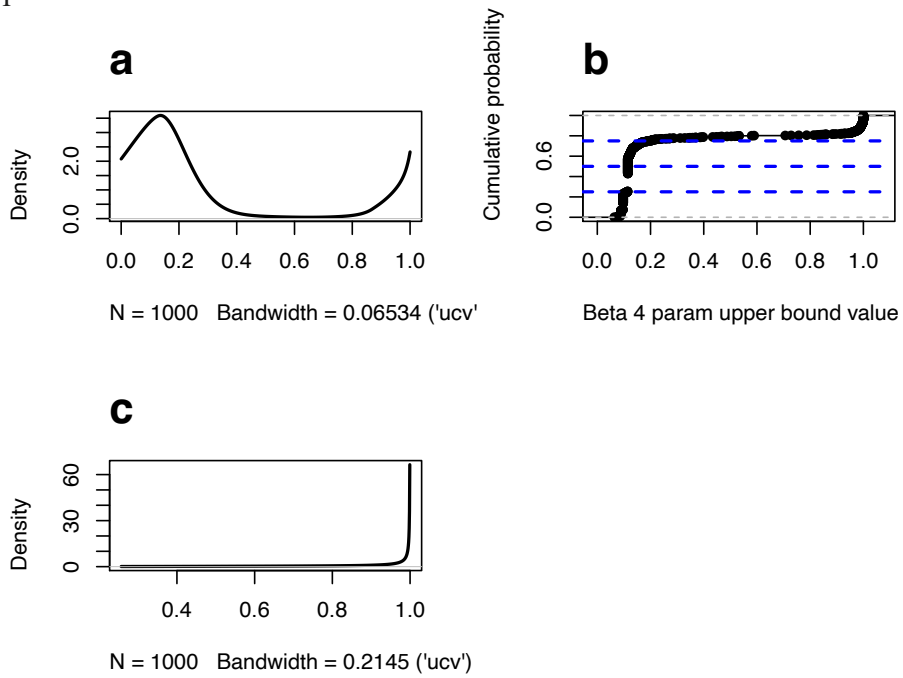

Figure 7. Probability density distribution of the upper bound of the 4 parameter beta for a)Argentina and c)Bangladesh, from 1,000 bootstrapped random draws with b) the cumulative density function for Argentina is added, with the 0.25, 0.5, and 0.75 quantiles marked in blue dotted lines.

4.2 Differences in insurance premiums across the probability distributions

We compare how insurance premiums in the insurance contract described in equation 1 vary depending on which distribution we choose to use. Recall that P is the payout required to cover losses, and the risk tolerance of the insurer is k. We ground this example by comparing premium estimates for likely index insurance design triggers and desired coverage given the fractional inundation and damage history in each location. Argentina purchases a policy to cover $100 million in losses (P=100) that triggers at 0.09 fractional flooded area, and Bangladesh is considering a policy to cover $200 million (P=200) that triggers at either 0.5 or 0.4 fractional flooded area (Figure 8a). We consider a sigmoid loss function per fractional flooded area (x) in equation 8, where y represents the approximate maximum amount of damage likely to occur given the flood history, z the slope of the relationship between damage and fractional flooded area, x^* the fractional inundated area at the inflection point, peak, or highest slope of the damage curve, and α the lower limit of recorded damage. We use reported damages in billions ($USD, see table 1) and inundated area history in each location to parametrize the loss function, assuming y=1.4 , z=80, and α=0.07 for Argentina and y=5, z=10, and α=0.7 for Bangladesh, and plot the estimated loss across a range of fractional flooded area (Figure 8a).

$$f(x) = \frac{y}{1+e^{-z\max(0,x-x^*)}} - \alpha \qquad (8)$$

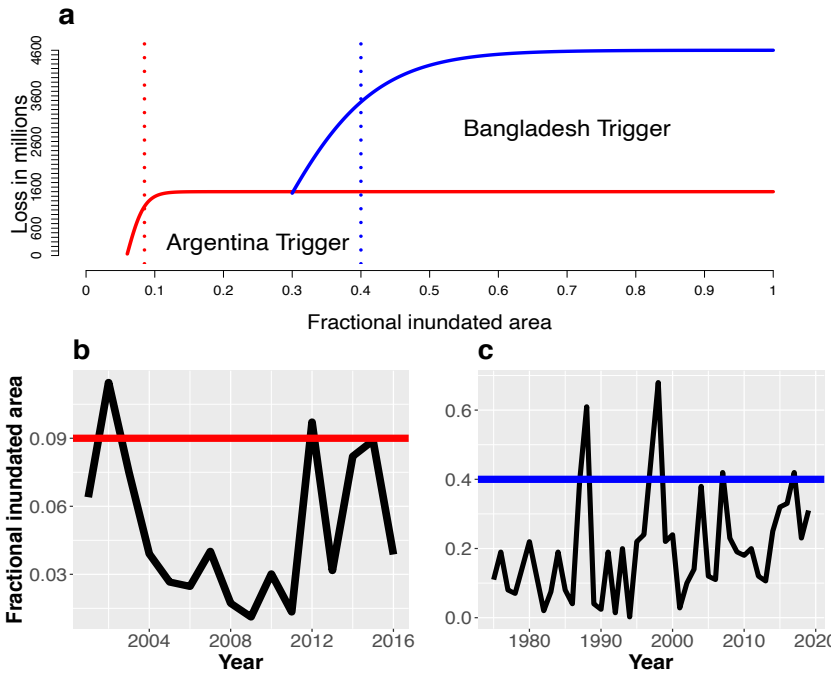

Figure 8. Proposed index insurance trigger and flood history. A) Estimated loss in millions per fractional inundated area and proposed index insurance trigger for Argentina (red) and Bangladesh (blue). B) Fractional inundated area in Argentina, with trigger in red. C) Fractional inundated area in Bangladesh, with trigger in blue.

We examine the cost of insurance premiums using equation 1 using the GEV, the 2-parameter Beta, and the 4 parameter Beta distribution for a policy chosen by each region (Figure 8).

Policyholders can choose at which fractional inundated area they want to trigger the policy and receive the payout. Other designs are possible, where coverage increases with larger fractional inundated area, but is held constant at one coverage in this example ($200 million for Bangladesh and $100 million for Argentina). The policyholder must consider at what fractional inundated area they expect to have a high loss (e.g. Figure 8a) that their capital reserves cannot cover. For example, the Indian government allocated ~$4 billion in USD for disaster response (Allocation and Release of Funds from the State Disaster Response Mitigation Fund during 2020-2021 2020), in 2020 and would only want to purchase a policy for extreme events at a fractional flood area that causes more damage than their available cover in the government budget.

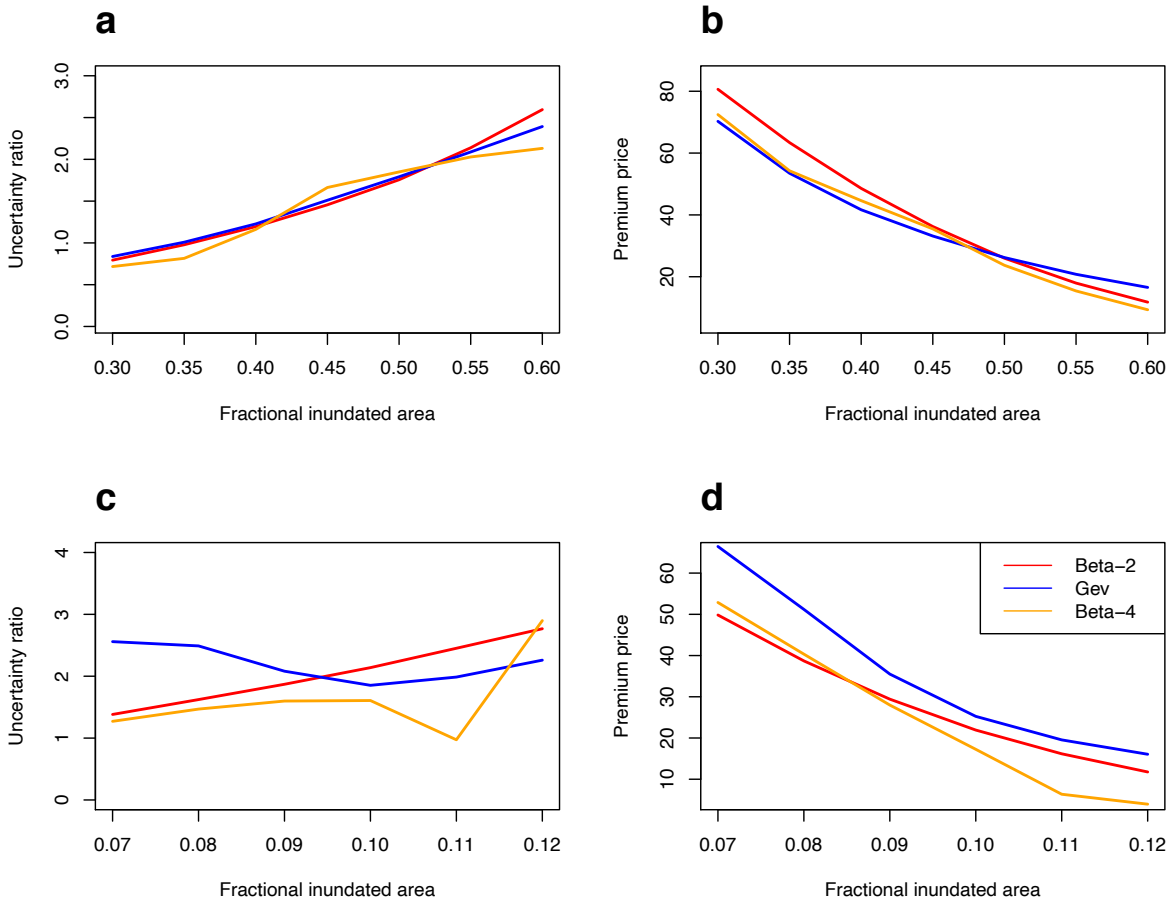

Figure 9. Estimated uncertainty ratio (risk premium/fair premium) (a and c for Bangladesh and Argentina and cost of premiums across fractional inundated area (b and d) for Bangladesh and Argentina for different probability distributions across fractional inundated area.

The risk premium is increasingly larger than the fair premium as fractional inundated area increases, as measured by the Uncertainty Ratio (risk premium/fair premium) (Figure 9a, c). In small sample sizes (which will be typical for writing insurance policies on a relatively short satellite record), the risk premium will nearly always be higher than the fair premium (Uncertainty Ratio >1). The risk premium decreases with larger sample sizes (e.g. Figure 3, comparing the uncertainty ratio for n=20 vs. n=200), and for smaller fractional flooded area. However, most policy holders will desire policies that both i) offer coverage at larger fractional

inundated areas that represent losses larger than their reserves and ii) keep insurance premium costs low while result in seeking coverage for extreme events at larger fractional flooded areas. Uncertainty plays a larger role in insurance premium prices with more extreme events.

Given policy holders will often desire payouts at larger return periods, we do not consider the GEV as a suitable probability distribution to price insurance premiums. The GEV can be dismissed on theoretical grounds because it is unbounded (and fractional flooded area can never be >1). Uncertainty estimates are wider using the GEV, which quickly include values >1 by the 50 or 100 year return period (Figure 5f, 6f). Figure 9 also demonstrates that the GEV results in a more expensive insurance premium at all fractional flooded areas in Argentina, and at fractional flooded areas >0.5 in Bangladesh.

Table 2. Estimated annual premium and flood losses per country, in units of millions of USD. Flood losses are estimated using the loss function in Figure 8, using the empirical time series in Figure 4. The estimated exceedance probability (and return periods) are in columns 3 and 4 for 2 Beta and 4 Beta estimates, respectively for the parameters considered. Column 5 is the number of times each trigger would have paid out based on the empirical flood history (Figure 8a red line, 8b blue line). Column 6 is the cumulative flood losses during this time period (1975-2019 in Bangladesh, 2001-2017 in Argentina) by multiplying the fractional flooded area per year with the associated estimated loss from equation 8 above the threshold/trigger (Figure 8a). The estimated annual premium for a trigger (0.09 for Argentina and 0.4 vs. 0.5 for Bangladesh) for the 2- and 4-parameter Beta are estimated in columns 7 and 8 (Table 2), along with the total premiums paid to the insurer in parentheses over the time period analyzed. The final two columns (9 and 10) are the ratio of total payouts to the region by the insurer (numbers of payouts times coverage (column 5) to the premiums paid by the government during the time period analyzed.

Country and years analyzed	Fractional flood trigger threshold	P_{exc} 2 Beta; return period	P_{exc} 4 Beta; return period	Times paid out (cover M USD)	Total flood losses M USD above trigger	2 Beta premium (total paid) M USD	4 Beta premium (total paid) M USD	2 Beta ratio of payouts to premium	4 Beta ratio of payouts to premium
Bangladesh (1975-2019)	0.4	0.11; 8.7 yrs	0.12; 8.19 yrs	4 (200)	1769.87	48.2 (2120.4)	47.7 (2098.8)	0.377	**0.381**
Bangladesh (1975-2019)	0.5	0.05; 19.9 yrs	0.05; 18.2 yrs	2 (200)	896.93	26 (1144)	27.6 (1214.4)	**0.350**	0.330
Argentina (2001-2017)	0.09	0.11; 9.32 yrs	0.24; 4.24 yrs	3 (100)	263.33	29 (464)	28.3 (452.8)	0.647	**0.663**

We compare insurance premiums prices across the 2- and 4-parameter Beta distribution for Bangladesh and Argentina (Table 2) since these are both admissible distributions, and the main question is whether the upper bound is smaller than 1 or 1. The societal optimum of the ratio of payouts to premiums is 1, where over time, the premium the buyer pays is closer to the average annual payout by the insurer. We therefore bold the numbers in columns 9 and 10 to indicate which insurance policy is preferred and closer to the ideal "fair" ratio. Note that the ratio of payouts to premiums will never equal 1, because the insurer needs to make a profit and cover

uncertainty, and the government values large sums of capital it needs to spend right after an event.

The 4 parameter Beta distribution is likely a better choice for insurance design for Argentina, and results in a cheaper premium and higher average annual payout to premium ratio. This is because inundation in Argentina is likely bounded far below 1 (the largest recorded fractional flood area is 0.12), and the additional parameters in the distribution help constrain the estimation (Figure 6). In Bangladesh, either the 2 or 4 Beta distribution could be used, but the cheaper insurance premium depends on the fractional flood area trigger chosen. In Bangladesh, the bound is already estimated to be close to 1 (Figure 5), and the effort of estimating two additional parameters for the 4 Beta distribution generates additional costs in the uncertainty ratio at larger fractional flooded areas (e.g. 0.5).

In general, uncertainty of estimation increases with return periods (larger, more infrequent floods) and as fractional flooded area grows in both examples (Figure 9, a and b). This implies that the proportion the estimation uncertainty contributes (the risk premium) increases. The lesson here is there is not ONE best distribution for estimating uncertainty for all flood index insurance situations. Our results suggest the 4-parameter Beta distribution is likely a good choice, but in some situations where the fractional flooded area bound is near 1, the 2-parameter Beta distribution may also be justified. An analysis such as the one presented here can help insurers and the insured appreciate the nuances and make an informed decision.

5. Discussion and Conclusion

This paper was motivated by a desire to directly explore the estimation of the annual exceedance probabilities for the fractional area inundated by flooding applied to index insurance. Using two settings as examples, we have demonstrated that a distribution, such as the Beta that has bounded support may in general be preferable to the GEV distribution that is commonly used for modeling extreme values. Results indicate that uncertainty is much higher for GEV compared to the Beta distribution for the relatively short record inundation time series analyzed in Argentina, or for extreme events in Bangladesh. GEV estimation becomes physically unfeasible (predicts fractional inundation greater than 1) at the 50 year return period in one dataset, showing the importance of a bounded distribution to estimate exceedance probabilities for extremely large inundated areas.

Our experiments showed estimating uncertainty associated with the resulting annual exceedance probabilities can vary by setting. We found the 4 parameter Beta distribution to yield lower uncertainty in Argentina, while in Bangladesh the 2 parameter Beta distribution is also acceptable, because an upper bound of 1 seemed plausible. We do not suggest that a particular distribution, or a fitting method for the parameters of a distribution be used prescriptively across all inundation risk analyses. The analyst should consider a strategy for each application that effectively reduces uncertainty. The choice of distribution should depend on two factors, i) the goodness of fit of that distribution to the data, as measured by an appropriately estimated log-likelihood of fit (or equivalent measure such as the BIC) and ii) the estimated uncertainty of the "predicted" distribution over the range of application of the distribution. In particular, we suggest assessing the uncertainty under extrapolation to the annual exceedance probabilities or return periods of interest when using short records such as satellite data. A more complex distribution that fits the observations better (factor i) may actually posit considerably higher or lower uncertainty under extrapolation to return periods amenable to index insurance applications (factor ii).

In this example, results show the distribution of the log-likelihood of the fit to the candidate distributions was often not distinguishable. The power of typical tests of goodness of fit of distributions is generally low with small sample sizes. What that means in practical terms is that the PPCC, Kolmogorov-Smirnov, or the Chi-Square tests used to choose between candidate probability distributions are a) not very robust in indicating which one is better, and b) that they are heavily weighted towards in sample performance, rather than the quality of fit in the tails of the distribution. This observation is what has often motivated the use of the Hill's (Hill 1975) and other tail index methods for extreme values. However, Moon et al. (1993) noted that such estimators typically have significantly higher uncertainty. The same argument applies to more complex distributions. Consequently, our argument in this paper is that for inundated area, which is anticipated to be bounded, the GEV that is unbounded would in many cases be conceptually inadmissible. However, for relatively small samples, the GEV may indeed exhibit a better fit in terms of log-likelihood than even a two-parameter Beta distribution. However, this may come with substantially higher uncertainty for the p_{exc} associated with the threshold selected for the index.

Parametric or index insurance policies would be written for the exceedance probability of a specific area inundated. A practical approach, beyond "fit" of the distribution, is to evaluate how the uncertainty ($sp_{exc)}$ around the exceedance probability ($p_{exc)}$ affects the insurance premium associated with that distribution for the desired inundated area trigger. Equation 1 can be used to estimate the premium price for a specified level of coverage, P, the needed payout to cover the damages associated with that level of inundation. The sensitivity of the premium to distribution choice covering both the expected value and the spread of p_{exc} for that threshold of inundated area can thus be evaluated to ultimately inform the choice of distribution. We emphasize that where distributions have a comparable fit to the data, the one which indicates a lower uncertainty is preferred, since it reflects a parsimony argument. However, we fully recognize that the true model uncertainty reflecting choice across distributions may be higher still, and this is reflected to an extent in the risk premium multiplier, k, that the insurer may choose. Given smaller sample sizes, the insurer may apply higher values of k reflecting their higher risk aversion. Of course this will magnify the impact of the uncertainty of estimation on the premium, and reinforce the choice of a more parsimonious model. The apparent duplicity in this reality is understood as the factorization of the insurer's risk perception (higher k) and the appropriateness of a model selection conditional on that as the one that is most consistent with the data in terms of fit and parsimony.

We focused on the variability of the predicted exceedance probability curve in the tails of the distribution, where a prospective insurance product is targeted. When using a short (<30 year) record to fit a distribution for a product that pays out at the 100-year return period, the performance or uncertainty on the extrapolated "tail" of the distribution matters most. For example, the fit of a 4-parameter distribution (where bounds are estimated instead of prescribed them as 0,1) may be indistinguishable from the 2-parameter distribution, and the p_{exc} corresponding to R_{x*} (the trigger) may be the same. However, if the uncertainty, or $sp_{exc,}$ would be higher for the 4-parameter distribution, then the 2 parameter distribution is a better choice, as in the Bangladesh example. On the other hand, if an upper bound of 1 is not justified, as indicated for Argentina, then the 4 parameter distribution does offer a better fit and lower uncertainty. If the fit of the two distributions were significantly different, then the computed R_{x*} for the same candidate *x*, would provide the basis for comparison across the two distributions,

since it would account for both the change in the p_{exc} and the sp_{exc}. As x is varied, the conclusion as to which one is better could very well change.

Our intent with these two examples was to emphasize that the choice of distribution and estimation procedure should be based on the target application for index insurance, instead of a prescriptive choice. Using only a goodness of fit measure fails to account for the typical extrapolation of the fit beyond the length of the record relevant for insurance or other catastrophic risk instrument design.

Focusing on uncertainty in index insurance design (over a goodness of fit measure) yields useful information to help governments arrive at an informed choice of which policy to purchase, with what coverage, and at what exceedance probability they desire payouts. Governments (and any insurance purchaser) have to make tradeoffs between the cost of investing in flood mitigation which can prevent frequent floods at high exceedance probabilities versus an insurance policy to cover more frequent events. Comparing the ratio of payouts to premiums (e.g. in table 2), compared to the relative costs of mitigation can aid that decision. Cheaper index insurance premiums to only cover rare catastrophic events at higher fractional flooded areas (Figure 9 b and d) can seem attractive, but implies the buyer is paying large portions of premium to cover the uncertainty associated with estimating exceedance probabilities of these rare events (Figure 9 a and c). Ultimately governments have to examine a holistic suite of flood adaptation options, and decide what type of insurance fits their risk mitigation strategy and budget best.

A major limitation of the proposed approach and important in future work is the issue of nonstationarity. Two major factors, climate change and flood adaptation efforts, change the relationship between fractional flood area and damage. Investing in flood mitigation significantly lowers the loss experienced as a function of inundated area (Jongman et al. 2015, Boulange et al. 2021). Table 1, for example, shows how fractional inundated area that was financially devastating in Bangladesh in the 70s and 80s (>4 billion), now cause less the 1 billion dollars of damage, likely due to extensive investments in embanks to protect infrastructure and agriculture (Rahman and Salehin 2013). This makes estimating the necessary coverage for fractional inundated area difficult. As floods increase in frequency and magnitude with a changing climate, exceedance probabilities with specific fractional inundated area could be underestimated using the past record. Yet current climate model projections for extreme event probabilities may be too uncertain to usefully inform insurance contracts at sub-regional resolutions (Fiedler et al. 2021). Both adaptation and climate change should be addressed in future work to improve index insurance design that reflects the needs of risk transfer products for the Anthropocene.

Index insurance, catastrophe bonds, and other types of risk transfer instruments could play an important role in adapting to floods and ensuring sustainable development. The direct use of inundation as a measure of flood risk, using satellite imagery for both risk estimation and verification of payoff, opens up an interesting set of questions for spatial analysis as well as insurance product design. The use of a direct measure could potentially reduce the uncertainties associated with the chain of hydrologic and climate models that are used for inundation estimates. On the other hand, the short records from satellites lead to higher uncertainty. The fact that the random variable of interest is bounded suggests an alternate choice of distribution beyond the GEV that constrains the uncertainty associated with extreme events and ultimately reduces the price of insurance premiums, making them more accessible. Bayesian approaches not explored in this paper could further reduce uncertainty in small samples of satellite data to improve flood index design. How best can the satellite data on inundation, land-use change, and recent climate be best used to update the changing probability of inundation and loss and inform

risk transfer products? This is an open question that we hope the flood risk community will consider to complement traditional discharge and flood model-based approaches to understanding risk.

Acknowledgements and Data

This work is undertaken as part of the Columbia World Project, ACToday, Columbia University in the City of New York. Additional funding was provided by NASA New Early Career Investigators Program Earth Sciences. 20-NIP20-0224. 2021-2024 Understanding flood risk in human altered landscapes from cities to farms: inferences from satellites and machine learning, PI Tellman. Data for fractional inundated area is available on HydroShare for Argentina and Bangladesh: https://www.hydroshare.org/resource/aafd32f20dbe4d09bddfd8cdd8fb6061/.

References

Aerts, J. P. M., S. Uhlemann-Elmer, D. Eilander, and P. J. Ward. 2020. Comparison of estimates of global flood models for flood hazard and exposed gross domestic product: a China case study. *Nat. Hazards Earth Syst. Sci.* 20(12):3245–3260.

Allen, G. H., X. Yang, J. Gardner, J. Holliman, C. H. David, and M. Ross. 2020. Timing of Landsat Overpasses Effectively Captures Flow Conditions of Large Rivers. *Remote Sensing* 12(9):1510.

Allocation and Release of Funds from the State Disaster Response Mitigation Fund during 2020-2021. 2020, March. . National Disaster Management Institute of India.

Amarnath, G. 2020, April 22. First satellite-based insurance trial in Bangladesh helps farmers recover from flooding. *CGIAR*.

Arnaud, P., C. Bouvier, L. Cisneros, and R. Dominguez. 2002. Influence of rainfall spatial variability on flood prediction. *Journal of Hydrology* 260(1–4):216–230.

Bardsley, E. 2018. Technical note: The Weibull distribution as an extreme value alternative for annual maxima. *Hydrology and Earth System Sciences Discussions*:1–9.

Bates, P. D. 2022. Flood Inundation Prediction. *Annual Review of Fluid Mechanics* 54(1):287–315.

Beard, L. R. 1960. Probability estimates based on small normal-distribution samples. *Journal of Geophysical Research* 65(7):2143–2148.

Beard, L. R. 1997. Estimating Flood Frequency and Average Annual Damage. *Journal of Water Resources Planning and Management* 123(2):84–88.

Benami, E., Z. Jin, M. R. Carter, A. Ghosh, R. J. Hijmans, A. Hobbs, B. Kenduiywo, and D. B. Lobell. 2021. Uniting remote sensing, crop modelling and economics for agricultural risk management. *Nature Reviews Earth & Environment*.

Bernhofen, M. V., C. Whyman, M. A. Trigg, P. A. Sleigh, A. M. Smith, C. C. Sampson, D. Yamazaki, P. J. Ward, R. Rudari, F. Pappenberger, F. Dottori, P. Salamon, and H. C. Winsemius. 2018. A first collective validation of global fluvial flood models for major floods in Nigeria and Mozambique. *Environmental Research Letters* 13(10):104007.

Botero, B. A., and F. Francés. 2010. Estimation of high return period flood quantiles using additional non-systematic information with upper bounded statistical models. *Hydrology and Earth System Sciences* 14(12):2617–2628.

Boulange, J., N. Hanasaki, D. Yamazaki, and Y. Pokhrel. 2021. Role of dams in reducing global flood exposure under climate change. *Nature Communications* 12(1):417.

Caribbean Catastrophe Risk Insurance Facility. 2015, March. Understanding CCRIF's Hurricane, Earthquake and Excess Rainfall Policies.

Chin, A. 2006. Urban transformation of river landscapes in a global context. *Geomorphology* 79(3):460–487.

Chowdhury, J. U., J. R. Stedinger, and L. -H Lu. 1991. Goodness-of-fit tests for regional generalized extreme value flood distributions. *Water Resources Research* 27(7):1765–1776.

Clement, K. Y., W. J. Wouter Botzen, R. Brouwer, and J. C. J. H. Aerts. 2018. A global review of the impact of basis risk on the functioning of and demand for index insurance. *International Journal of Disaster Risk Reduction* 28:845–853.

Douglas, E. M., and R. M. Vogel. 2006. Probabilistic Behavior of Floods of Record in the United States. *J. Hydrol. Eng.* 11(6)(December):631–635.

E. Vermote and R. Wolfe. 2015. MOD09GQ MODIS/Terra Surface Reflectance Daily L2G Global 250m SIN Grid V006. NASA EOSDIS LP DAAC.

Enenkel, M., D. Osgood, M. Anderson, B. Powell, J. McCarty, C. Neigh, M. Carroll, M. Wooten, G. Husak, C. Hain, and M. Brown. 2019. Exploiting the Convergence of Evidence in Satellite Data for Advanced Weather Index Insurance Design. *Weather, Climate, and Society* 11(1):65–93.

Fiedler, T., A. J. Pitman, K. Mackenzie, N. Wood, C. Jakob, and S. E. Perkins-Kirkpatrick. 2021. Business risk and the emergence of climate analytics. *Nature Climate Change*.

Finer, M., S. Novoa, M. J. Weisse, R. Petersen, J. Mascaro, T. Souto, F. Stearns, and R. G. Martinez. 2018. Combating deforestation: From satellite to intervention. *Science* 360(6395):1303–1305.

Gilleland, E., and R. W. Katz. 2016. ExtRemes 2.0: An extreme value analysis package in R. *Journal of Statistical Software* 72(8).

Gleason, C. J., and M. T. Durand. 2020. Remote Sensing of River Discharge: A Review and a Framing for the Discipline. *Remote Sensing* 12(7):1107.

Gleason, C. J., and A. N. Hamdan. 2017. Crossing the (watershed) divide: satellite data and the changing politics of international river basins. *Geographical Journal* 183(1):2–15.

Grill, G., B. Lehner, M. Thieme, B. Geenen, D. Tickner, F. Antonelli, S. Babu, P. Borrelli, L. Cheng, H. Crochetiere, H. Ehalt Macedo, R. Filgueiras, M. Goichot, J. Higgins, Z. Hogan, B. Lip, M. E. McClain, J. Meng, M. Mulligan, C. Nilsson, J. D. Olden, J. J. Opperman, P. Petry, C. Reidy Liermann, L. Sáenz, S. Salinas-Rodríguez, P. Schelle, R. J. P. Schmitt, J. Snider, F. Tan, K. Tockner, P. H. Valdujo, A. van Soesbergen, and C. Zarfl. 2019. Mapping the world's free-flowing rivers. *Nature* 569(7755):215–221.

Hannah, D. M., S. Demuth, H. A. J. van Lanen, U. Looser, C. Prudhomme, G. Rees, K. Stahl, and L. M. Tallaksen. 2011. Large-scale river flow archives: Importance, current status and future needs. *Hydrological Processes* 25(7):1191–1200.

Haraguchi, M. 2018. Innovations towards Climate-Induced Disaster Risk Assessment and Response. Columbia.

Hawker, L., J. Neal, B. Tellman, J. Liang, G. Schumann, C. Doyle, J. A. Sullivan, J. Savage, and R. Tshimanga. 2020. Comparing earth observation and inundation models to map flood hazards. *Environmental Research Letters*.

Hill, B. M. 1975. A Simple General Approach to Inference About the Tail of a Distribution. *The Annals of Statistics* 3(5).

Hultquist, C., and G. Cervone. 2020. Integration of Crowdsourced Images, USGS Networks, Remote Sensing, and a Model to Assess Flood Depth during Hurricane Florence. *Remote Sensing* 12(5):834.

Ibarlucía, D., F. Carmona, C. Mancino, M. Bayala, M. Silicani, F. Degano, R. Rivas, G. Cazenave, M. Varni, V. Barbero, P. Toscani, I. Castany, D. Ramírez, L. Aguirre, F. Oyarbide, and A. Ramallo. 2017. RED DE MONITOREO DE EVENTOS HIDROLÓGICOS EXTREMOS EN LA VERTIENTE SUR DEL RÍO SALADO, PROVINCIA DE BUENOS AIRES:14.

Islam, A. S., A. Haque, and S. K. Bala. 2010. Hydrologic characteristics of floods in Ganges–Brahmaputra–Meghna (GBM) delta. *Natural Hazards* 54(3):797–811.

Jensen, K., and K. Mcdonald. 2019. Surface Water Microwave Product Series Version 3: A Near-Real Time and 25-Year Historical Global Inundated Area Fraction Time Series From Active and Passive Microwave Remote Sensing. *IEEE Geoscience and Remote Sensing Letters* 16(9):1402–1406.

Jongman, B., H. C. Winsemius, J. C. J. H. Aerts, E. Coughlan de Perez, M. K. van Aalst, W. Kron, and P. J. Ward. 2015. Declining vulnerability to river floods and the global benefits of adaptation. *Proceedings of the National Academy of Sciences*:201414439.

Khalil, A. F., H.-H. Kwon, U. Lall, M. J. Miranda, and J. Skees. 2007. El Niño-Southern Oscillation-based index insurance for floods: Statistical risk analyses and application to Peru: ENSO-BASED INSURANCE INDEX FOR FLOODS. *Water Resources Research* 43(10).

Klein, I., A. Dietz, U. Gessner, S. Dech, and C. Kuenzer. 2015. Results of the Global WaterPack: a novel product to assess inland water body dynamics on a daily basis. *Remote Sensing Letters* 6(1):78–87.

Kuenzer, C., I. Klein, T. Ullmann, E. Georgiou, R. Baumhauer, and S. Dech. 2015. Remote Sensing of River Delta Inundation: Exploiting the Potential of Coarse Spatial Resolution, Temporally-Dense MODIS Time Series. *Remote Sensing* 7(7):8516–8542.

Latrubesse, E. M., and D. Brea. 2009. Floods in Argentina. Pages 333–349 *Developments in Earth Surface Processes*. Elsevier.

Liang, J., and D. Liu. 2020. Estimating Daily Inundation Probability Using Remote Sensing, Riverine Flood, and Storm Surge Models: A Case of Hurricane Harvey. *Remote Sensing* 12(9):1495.

Liu, X., Y. Huang, X. Xu, X. Li, X. Li, P. Ciais, P. Lin, K. Gong, A. D. Ziegler, A. Chen, P. Gong, J. Chen, G. Hu, Y. Chen, S. Wang, Q. Wu, K. Huang, L. Estes, and Z. Zeng. 2020. High-spatiotemporal-resolution mapping of global urban change from 1985 to 2015. *Nature Sustainability*.

Martins, S., and J. R. Stedinger. 2000. Generalized maximum likelihood generalized exteme value quantile estimators for hydrologic data. *Water Resources Research* 36(3):737–744.

Matheswaran, K., N. Alahacoon, R. Pandey, and G. Amarnath. 2019. Flood risk assessment in South Asia to prioritize flood index insurance applications in Bihar, India. *Geomatics, Natural Hazards and Risk* 10(1):26–48.

Merwade, V., F. Olivera, M. Arabi, and S. Edleman. 2008. Uncertainty in Flood Inundation Mapping: Current Issues and Future Directions. *Journal of Hydrologic Engineering* 13(7):608–620.

Merz, B., H. Kreibich, and U. Lall. 2013. Multi-variate flood damage assessment: a tree-based data-mining approach. *Natural Hazards and Earth System Sciences* 13(1):53–64.

Merz, B., H. Kreibich, R. Schwarze, and A. Thieken. 2010. Review article "Assessment of economic flood damage." *Natural Hazards and Earth System Science* 10(8):1697–1724.

Monirul Qader Mirza, M. 2002. Global warming and changes in the probability of occurrence of floods in Bangladesh and implications. *Global Environmental Change* 12(2):127–138.

Moon, Y.-I., U. Lall, and K. Bosworth. 1993. A comparison of tail probability estimators for flood frequency analysis. *Journal of Hydrology* 151(2–4):343–363.

Norton, M. T., C. Turvey, and D. Osgood. 2012. Quantifying spatial basis risk for weather index insurance. *The Journal of Risk Finance* 14(1):20–34.

Osgood, D., B. Powell, R. Diro, C. Farah, M. Enenkel, M. Brown, G. Husak, S. Blakeley, L. Hoffman, and J. McCarty. 2018. Farmer Perception, Recollection, and Remote Sensing in Weather Index Insurance: An Ethiopia Case Study. *Remote Sensing* 10(12):1887.

Papadakis, M., M. Tsagris, M. Dimitriadis, S. Fafalios, I. Tsamardinos, M. Fasiolo, G. Borboudakis, J. Burkardt, C. Zou, K. Lakiotaki, and C. Chatzipantsiou. 2020. *Rfast: A Collection of Efficient and Extremely Fast R Functions*.

Pohlert, T. 2020. Package ' ppcc ' 1.

Policelli, F., D. Slayback, B. Brakenridge, J. Nigro, A. Hubbard, B. Zaitchik, M. Carroll, and H. Jung. 2016. The NASA Global Flood Mapping System. Page 47 *Remote Sensing of Hydrological Extremes*. Springer.

Rahman, R., and M. Salehin. 2013. Flood risks and reduction approaches in Bangladesh. Pages 65–90 *Disaster risk reduction approaches in Bangladesh*. Springer.

Razali, N. M., and Y. B. Wah. 2011. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics* 2(1):21–33.

Rivas, R. 2017, May 17. Mediciones ambientales para la previsión de eventos futuros. Instituto de Hidrología de Llanuras.

Scotti, V., M. Giannini, and F. Cioffi. 2020. Enhanced flood mapping using synthetic aperture radar (SAR) images, hydraulic modelling, and social media: A case study of Hurricane Harvey (Houston, TX). *Journal of Flood Risk Management* 13(4).

Sebastian, A., A. Gori, R. B. Blessing, K. van der Wiel, and B. Bass. 2019. Disentangling the impacts of human and environmental change on catchment response during Hurricane Harvey. *Environmental Research Letters* 14(12):124023.

Shastry, A., M. Durand, J. Neal, A. Fernández, S. C. Phang, B. Mohr, H. C. Jung, S. Kari, M. Moritz, B. G. Mark, S. Laborde, A. Murumkar, and I. Hamilton. 2020. Small-scale anthropogenic changes impact floodplain hydraulics: Simulating the effects of fish canals on the Logone floodplain. *Journal of Hydrology* 588:125035.

Smith, J. A., A. A. Cox, M. L. Baeck, L. Yang, and P. Bates. 2018. Strange Floods: The Upper Tail of Flood Peaks in the United States. *Water Resources Research* 54(9):6510–6542.

Smith, M. B., V. I. Koren, Z. Zhang, S. M. Reed, J.-J. Pan, and F. Moreda. 2004. Runoff response to spatial variability in precipitation: an analysis of observed data. *Journal of Hydrology* 298(1–4):267–286.

Sofia, G., and E. I. Nikolopoulos. 2020. Floods and rivers: a circular causality perspective. *Scientific Reports* 10(1):1–17.

Stedinger, J. R. 1983. Design events with specified flood risk. *Water Resources Research* 19(2):511–522.

Stedinger, J. R. 1997. Expected Probability and Annual Damage Estimators. *Journal of Water Resources Planning and Management* 123(2):125–135.

Surminski, S., L. M. Bouwer, and J. Linnerooth-Bayer. 2016. How insurance can support climate resilience. *Nature Climate Change* 6(4):333.

Surminski, S., and D. Oramas-Dorta. 2014. Flood insurance schemes and climate adaptation in developing countries. *International Journal of Disaster Risk Reduction* 7:154–164.

Swain, D. L., O. E. J. Wing, P. D. Bates, J. M. Done, K. A. Johnson, and D. R. Cameron. 2020. Increased Flood Exposure Due to Climate Change and Population Growth in the United States. *Earth's Future* 8(11).

Tai, K. C. 1987. Flood Risk Bias Analysed through a Multi-State Flood Insurance Model. Pages 395–405 *in* V. P. Singh, editor. *Application of Frequency and Risk in Water Resources: Proceedings of the International Symposium on Flood Frequency and Risk Analyses, 14–17 May 1986, Louisiana State University, Baton Rouge, U.S.A*. Springer Netherlands, Dordrecht.

Tellman, B., J. A. Sullivan, C. Kuhn, A. J. Kettner, C. S. Doyle, G. R. Brakenridge, T. A. Erickson, and D. A. Slayback. 2021a. Satellite imaging reveals increased proportion of population exposed to floods. *Nature* 596(7870):80–86.

Tellman, B., J. Sullivan, and C. Doyle. 2021b. Global Flood Observation with Multiple Satellites: Applications in Rio Salado, Argentina, and the Eastern Nile Basin. Page 352 *in* H. Yu, D. P. Lettenmaier, T. Quihong, and P. J. Ward, editors. *Global Drought and Flood: Monitoring, Prediction, and Adaptation*. Wiley.

Teng, J., A. J. Jakeman, J. Vaze, B. F. W. Croke, D. Dutta, and S. Kim. 2017. Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. *Environmental Modelling & Software* 90:201–216.

Trigg, M. A., C. E. Birch, J. C. Neal, P. D. Bates, A. Smith, C. C. Sampson, D. Yamazaki, Y. Hirabayashi, F. Pappenberger, E. Dutra, P. J. Ward, H. C. Winsemius, P. Salamon, F. Dottori, R. Rudari, M. S. Kappes, A. L. Simpson, G. Hadzilacos, and T. J. Fewtrell. 2016. The credibility challenge for global fluvial flood risk analysis. *Environmental Research Letters* 11(9):094014.

USGS. 1982. Guidelines for determining flood flow frequency. US Deprtment of the Interior.

Vijay P. Singh, Ph. D., D. Sc. ,. D. Eng. (Hon.), Ph. D. (Hon.), D. Sc. (Hon.), P. E. ,. P. H. ,. Hon. D. WRE, Academician (GFA). 2017. *Handbook of Applied Hydrology, Second Edition*. 2nd edition. McGraw-Hill Education, New York.

Vogel, R. M. 1986. The Probability Plot Correlation Coefficient Test for the Normal, Lognormal, and Gumbel Distributional Hypotheses. *Water Resources Research* 22(4):587–590.

Wang, J. Z. 2005. A Note on Estimation in the Four-Parameter Beta Distribution. *Communications in Statistics - Simulation and Computation* 34(3):495–501.

Ward, P. J., B. Jongman, P. Salamon, A. Simpson, P. Bates, T. De Groeve, S. Muis, E. C. de Perez, R. Rudari, M. a. Trigg, and H. C. Winsemius. 2015. Usefulness and limitations of global flood risk models. *Nature Climate Change* 5(8):712–715.

Wing, O. E. J., P. D. Bates, C. C. Sampson, A. M. Smith, K. A. Johnson, and T. A. Erickson. 2017. Validation of a 30 m resolution flood hazard model of the conterminous United States: 30 m RESOLUTION FLOOD MODEL OF CONUS. *Water Resources Research* 53(9):7968–7986.

World Food Program. 2017. Advocacy paper Monsoon 2017. Bangladesh Food Security Cluster.

Wu, Haizhen, Godfrey, Jonathan R, and Govin-Daraju, Kondaswamy. 2020. *ExtDist*.
Yee, T. W. 2010. Journal of Statistical Software The VGAM Package for Categorical Data Analysis 32(10).
Zajic, B. 2019, June 17. How Flood Mapping From Space Protects The Vulnerable And Can Save Lives. *Planet Blog*.