Counterfactual Risk Analysis of Tropical Cyclones

Master Thesis
Oeschger Centre for Climate Change Research
Faculty of Science, University of Bern

handed in by
Tamara Alessandra Baumann

2022

Supervisor:
Prof. Dr. Olivia Romppainen-Martius

Advisor:
Dr. Alessio Ciullo

Abstract

Tropical cyclones (TC) are among the most devastating natural hazards causing losses and damages in almost all tropical regions. However, an accurate risk analysis of TC in these regions is not straightforward due to sparse data availability and short observational records. High impact extreme events at the tail of the distribution are often underestimated and biased towards the outcome of past events. The goal of counterfactual risk analysis is to improve risk assessment of extreme events by incorporating data on near-miss events. Considering the chaotic nature of atmospheric dynamics, the outcome of past TC can be viewed as only one of many possible realizations. By incorporating alternative but physically plausible scenarios of past events, so called counterfactuals, the sparse observational datasets of TC can be improved significantly.

Within this thesis, TC forecasts of past events are used as a source to create counterfactual scenarios of past yearly landfall rates. By identifying possible landfalls of observed events and worst-case scenarios, downward counterfactuals are established and compared to observed landfall rates using Bayesian inference.

The analysis of downward counterfactual scenarios leads to a higher estimation of expected mean yearly landfall rates, probability of extreme events and outliers, and future annual landfall rates, including its extremes. TC forecasts allow the detection of near-miss events and expansion of the sparse observational data, especially for small island states where TC landfalls are rare. The results highlight that counterfactual data is promising in improving the risk assessment of TC at the tail of the distribution. The results also show that TC forecasts can be a great source in building counterfactual scenarios.

Abbreviations

Variables

$ymax$ ....highest observed yearly landfall rate up to 2007

μ_{hist} ....mean observed landfall rate up to 2007

μ_{hist50} .... mean observed landfall rate 1950-2007

μ_{hist80} .... mean observed landfall rate 1980-2007

$p(\lambda_{50})$.. prior probability based on observation period 1950-2007

$p(\lambda_{80})$.. prior probability based on observation period 1980-2007

$p(\lambda|y_{obs})$..posterior probability based on observed y

$E[\lambda|y_{obs}]$ .... expected value for λ / mean of posterior distribution based on observed y

$p(\lambda|y_{CFmax})$ .... posterior probability based on

$E[\lambda|y_{CFmax}]$ ....expected value for λ / mean of posterior distribution based on worst-case

$\Delta E[\lambda|y]$..$E[\lambda|y_{CFmax}] - E[\lambda|y_{obs}]$

$\tilde{y}$..future landfall rates

$p(\tilde{y}|y_{obs})$..posterior predictive probability based on observed y

$p(\tilde{y}|y_{CFmax})$.. posterior predictive probability based on counterfactual worst-case scenario for y

$p(\tilde{y} \geq ymax|y_{obs})$ .... probability future landfall rates are equal of higher than $ymax$ for PP based on observed y

$p(\tilde{y} \geq ymax|y_{CFmax})$.. probability future landfall rates are equal of higher than $ymax$ for PP based on counterfactual worst-case scenario for y

$\Delta p(\tilde{y} \geq ymax)$.. $p(\tilde{y} \geq ymax|y_{CFmax}) - p(\tilde{y} \geq ymax|y_{obs})$

$E[\tilde{y}_{95p}|y_{obs}]$ .... expected landfall rate at 95th percentile of PP distribution, observed

$E[\tilde{y}_{95p}|y_{CFmax}]$........................expected landfall rate at 95th percentile of PP distribution, counterfactual worst-case scenario for y

Contents

1 Introduction

Tropical cyclones (TC) are among the most devastating and deadliest natural hazards, causing enormous damage and losses in almost all tropical regions. One of the most prominent examples was Hurricane Katrina which made landfall (LF) on the US Gulf Coast close to New Orleans in late August 2005. The accompanying rainfalls and storm surge caused by Katrina lead to the flooding of 80 percent of the city of New Orleans. Overall, Katrina claimed 1800 lives and with damages around $160 billion was the costliest natural disaster in US history (Britannica, 2021). A new record in overall damages due to TCs was reached in 2017 when hurricanes Harvey, Irma and Maria made landfall within a span of only four weeks leading to overall losses around $220 billion (Munich RE, 2017). With their large impacts upon landfall, it is important to have a good understanding of the disaster risk associated with tropical cyclones. This knowledge can help decision makers to take necessary mitigation measures. Furthermore, (re)insurance companies are interested in reliable probability distributions to build their portfolios correctly.

Tropical cyclones are relatively rare with globally around 90 storms in a year. With reliable data only available for the last few decades the risk assessment of TCs is challenging (Bloemendaal *et al.*, 2020). Especially for extreme impact events with high return periods and low probabilities the available data is sparse and therefore complicating reliable risk assessment. When modelling disaster risk based solely on past events, extreme events can easily be underestimated, due to their low frequency of occurrence. Information on such extreme events is crucial, especially when assessing the vulnerabilities of communities and planning necessary mitigation measures. Therein, lies the danger of catastrophes occurring that were not within the modelled risk horizon (Woo, 2019). To better attribute for the described outcome bias, Gordon Woo in 2016 first introduced the concept of using downward counterfactual (CF) risk analysis to detect possible high impact events. By incorporating information on near miss events the risk assessment at the tail of the distribution can be improved (Woo, 2016).

The term counterfactual is more commonly used in cognitive psychology, which defines a downward counterfactual as a "thought about the past where the outcome was worse than what actually happened" (Woo, 2019). Woo (2019) suggests that instead of treating history as fixed, we can view it as one of many possible unfolding of past events. More common in risk analysis are upward counterfactuals, meaning thoughts about past events imagining a better outcome. This comes into play, for example, when analyzing how a past event could have been prevented or how its damages and losses could have been minimized.

In order to detect possible high impact events by using counterfactual risk analysis, researchers may consider how past events could have been worse if the circumstances were only slightly different (Woo, 2019). As such, counterfactuals are defined as physically plausible unfolding of past events illustrating a different but possible outcome. For tropical cyclones, this could mean analyzing how a past hazard event could possibly have developed considering, for example, higher wind speed on landfall, more precipitation, or stronger storm surges. Very important for the risk analysis of tropical cyclones is also their trajectory. Only slight alterations can determine whether a TC makes landfall or not or whether a community or city is hit by a TC or not.

As an example, one can look at Hurricane Ivan, which headed towards New Orleans in 2004 as a category 4 TC with wind speed of 225 km h^{-1} (Woo, Maynard and Seria, 2017). Luckily, the hurricane turned east and did not directly hit New Orleans, making landfall instead east of Mobile Bay, Alabama. Despite this given outcome previous forecast data of Hurricane Ivan showed the storm possibly striking New Orleans. Considering this preceding forecast data, the stochastic models indicated an entirely different loss distribution due to changes in the track geometry, the hurricane intensity on landfall, the height of the storm surge, and the consequent inland flood potential. The tail of this counterfactual loss distribution would have included the loss realized one year later by Hurricane Katrina.

This example of Hurricane Ivan already illustrates the potential of forecasting data as a possible source for counterfactual in risk analysis of TCs. Considering the chaotic nature of atmospheric dynamics only slight perturbations can lead to a completely different path and evolution of the storm. It is these small perturbations who decide which of the forecasted paths it follows closest to. However, at some point in time each forecasted path can be viewed as a possible outcome of the event. In this sense forecasting data of past events illustrate alternative, but physically plausible outcomes. As such, forecasting data is a good source for counterfactual information to expand the sparse observational data of TCs.

Building on the proposition of Woo (2016) there have been a few studies implementing downward counterfactuals in their risk analysis (e.g., Aspinall and Woo, 2019; Oughton *et al.*, 2019). Up to date, no studies specifically used counterfactual data in the risk analysis of tropical cyclones. Considering the challenges in TC risk analysis leading to an outcome bias towards past events, and a possible underestimation at the tail of the hazard risk distribution, including counterfactual data might be a powerful source in expanding the horizon of possible extreme events. The aim of this thesis is therefore to explore forecast data as source for counterfactual information of past TCs to expand the sparse observational data on which current risk analysis is based. Thus, the data basis for risk analysis can be augmented significantly. This is illustrated for the North Atlantic in Figure 1.1, where the density and amount of forecast tracks is much higher than by merely considering observed tracks.

More specific, this thesis compares past yearly landfall rates, or landfall counts, of TCs to counterfactual landfall rates based upon the combination of different forecast tracks of past TC events. By specifically looking for downward counterfactuals the goal is to detect possible extreme landfall rates. High or extreme landfall rates mark very stormy TC seasons in which the cumulated damages and losses might be hazardous. This illustrated the year 2017 with three landfalling TCs leading to the costliest year in TC damages. Thus, an exploration of how many landfalls within a year could be expected when including downward counterfactuals in the risk analysis can be of high importance for decision makers both regarding mitigation measures as well as in the (re)insurance sector.

The dataset on forecast data used in this thesis contains global TC data from 2008 to 2019 with several forecasts for each TC. Each forecast contains several forecast ensemble members, representing an alternative path to the observed TC. Considering all TCs in this time period, each combination of the individual forecast members corresponding to observed TCs can be

viewed as a counterfactual history or a counterfactual scenario. The resulting landfall rates of said counterfactual scenario are then the basis to assess change in expected landfall rates.
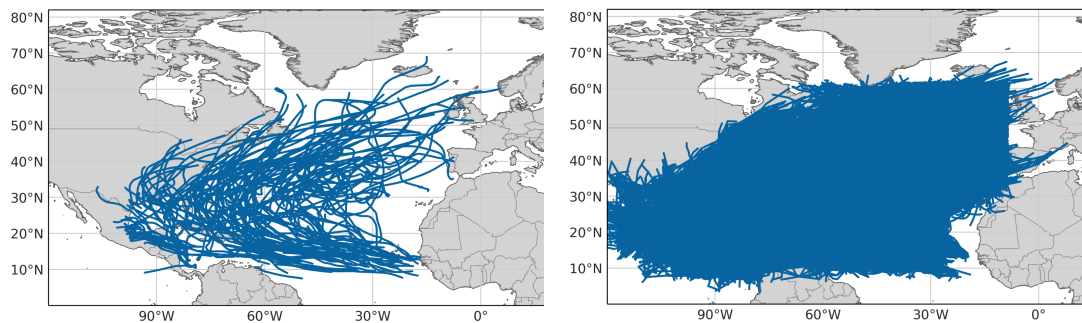

Figure 1.1: All observed TC tracks in the North Atlantic for 2008-2019 (left) compared to all forecast tracks of the same period (right).

Expected mean landfall rates, future landfall rates and extreme landfall rates based on counterfactuals are analyzed using Bayesian inference (BI). The advantage of Bayesian inference compared to the traditional frequentist statistical inference is that it incorporates the notion of learning from new information. As such previous knowledge or data on TC landfall rates can be included updated using observed or counterfactual landfall rates. Bayesian inference then allows to make projections for future landfall rates. Also, Bayesian inference returns probabilities that incorporate the uncertainty associated with the unknown variable of interest (Gelman *et al.*, 2020). As such, the probability of all possible values for mean or extreme landfall rates can be explored and compared between counterfactual and observed scenarios.

The previous knowledge of TC landfall rates incorporated in the analysis is based on observed yearly landfall rates up to 2007. This represents the period before forecast data as counterfactuals are available. The sensitivity of the BI towards this previous or prior information is tested by comparing different periods of historic landfall. Thus, the period 1950 – 2007 is compared to the period 1980 – 2007. Further, the sensitivity of the BI is also tested towards the used counterfactual information. The range of counterfactual scenarios is compared across forecast data from two different providers and a combination of both.

The Bayesian inference analysis was conducted separately for six ocean basins North Atlantic (NA), East Pacific (EP), West Pacific (WP), North Indian Ocean (NI), South Indian Ocean (SI), and South Pacific (SP) (Figure 1.2). It was conducted for the ocean basins as well for the individual countries within the respective basin.

When assuming a possible underestimation of extreme hazard events, it is of special interest to analyze whether future projections of TC tracks based on counterfactuals are more extreme, than historic data suggests. So, while several counterfactual scenarios are created and included in the analysis the focus of the thesis are downward counterfactual scenarios and how they change the risk assessment of yearly landfall rates of TCs. Overall, the following, interlinked, research questions are answered:

- Can forecast tracks be used as a source for counterfactual information in risk analysis of tropical cyclones?

- To what extent does accounting for downward counterfactual scenarios change the probability distribution of yearly landfall rates and future projections of yearly landfall?

- What are differences and similarities of findings across case studies for different
  … periods for prior information on landfall rates?
  … forecast dataset based on providing agency?
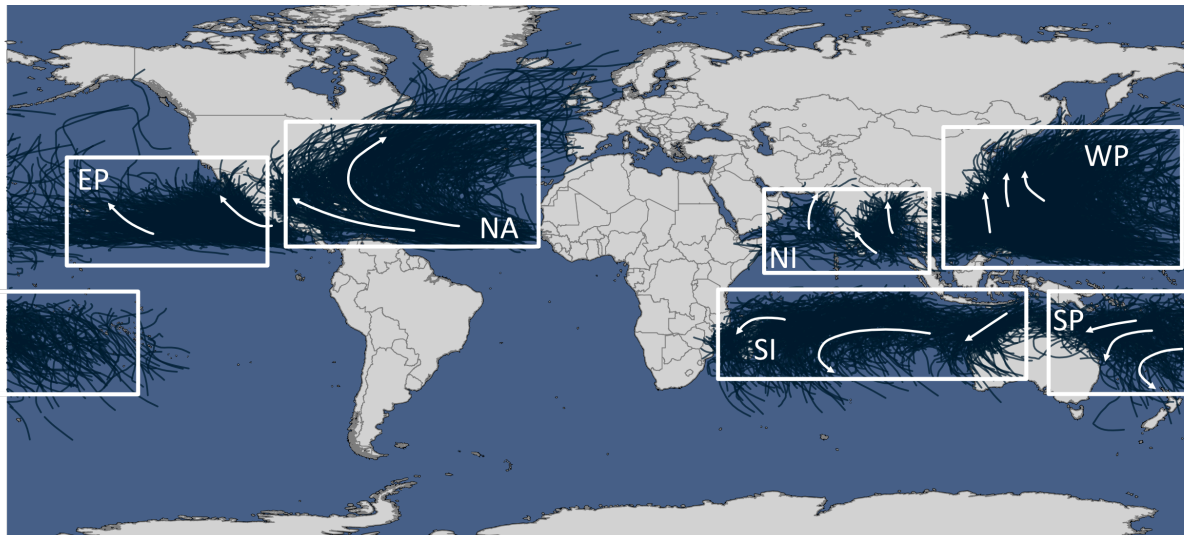 … basins and countries?

Figure 1.2: Overview tropical cyclone basins and observed TC tracks 1980-2019. Arrows indicate general movement. NA: North Atlantic, EP: East Pacific, WP: West Pacific, NI: North Indian Ocean, SI: South Indian Ocean, SP: South Pacific.

First in a background chapter risk of tropical cyclones and forecast data as source for counterfactual risk analysis of tropical cyclones are explored more thoroughly. After, the application of Bayesian inference within this thesis is presented, followed by a description of the data. In Chapter 4 the results for the different case studies are presented, compared, and discussed. Finally, in Chapter 5 follows the conclusion of the analysis and an outlook for counterfactual risk analysis of tropical cyclones.

2 Background

2.1 Tropical Cyclone Risk

Tropical cyclones are symmetric low-pressure systems that develop in the tropics and can reach sizes between 100 to 4000 km (Lohmann, Lüönd and Mahrt, 2016). They are characterized by a mostly cloudless eye of a diameter between 8 and 200 km. The eye is surrounded with the eyewall where convection and thunderstorm activity are strongest and underneath which the highest near surface wind speeds can be found. This is the most dangerous and destructive part of the TC. Spiraling outward from the eye there are often secondary cells arranged in bands. They are called rainbands. The major basins of tropical cyclone formation are the North Atlantic, East Pacific, West Pacific, North Indian Ocean, South Indian Ocean, and South Pacific (Figure 1.2). The largest and the most TCs occur in the West Pacific. Depending on their region of occurrence TCs are also called hurricanes, typhoons, severe tropical cyclones, tropical cyclones, or severe cyclonic storms.

The main driver of TCs is latent heat release as air driven over warm waters humifies, rises and the water condenses. When approaching land, they often intensify, due to the warm shallow water, reaching maximum intensity at landfall. Cut off from their main energy source they will weaken and dissipate over land also due to added friction at the surface.

The path a storm takes is referred to as the TC track. It marks the location of the center of the storm, or the point of lowest pressure (central pressure) at different time steps. Landfall occurs when all or parts of the eyewall passes across the coast. Accounting for the radius of the eyewall within this thesis landfall is assumed when the track crosses the coastline or comes within 50 km of the coast.

TCs are categorized by their intensity measured as maximum sustained wind speed over 1 minute at 10 meters above ground. Most often used is the Saffir-Simpson scale (Table 2.1) (Lohmann, Lüönd and Mahrt, 2016). The most severe category 5 refers to TCs with wind speeds exceeding 250 km h^{-1}. For extremely severe TCs maximum sustained winds above 300 km h^{-1} can be measured. When a storm develops the intensity of a tropical storm (sustained wind speed higher than 33 kt) it is given a name to distinguish it easily from other cyclonic systems.

Table 2.1: Categorization of tropical cyclone. Based on Saffir-Simpson scale (adapted from NOAA, National Hurricane Center and Central Pacific Hurricane Center, 2021; Wikipedia, 2021). Storm surge height and minimum pressure are for reference only. Categorization is based on wind speed.

Name	Category	Sustained winds	Storm surge (m)	Minimum pressure (hPa)	
Tropical Depression	-1	≤ 33 kt (≤ 62 km h-1)	0		
Tropical Storm	0	34-63 kt (63-118 km h-1)	0-0.9		Named storms
Tropical Cyclone	1	64-82 kt (119-153 km h-1)	1.0-1.7	980-994	
Tropical Cyclone	2	83-95 kt (154-177 km h-1)	1.8-2.6	965-979	
(Major) Tropical Cyclone	3	96-112 kt (178-208 km h-1)	2.7-3.8	945-964	
(Major) Tropical Cyclone	4	113-136 kt (209-251 km h-1)	3.9-5.6	920-944	
(Major) Tropical Cyclone	5	≥ 137 kt (≥ 252 km h-1)	≥ 5.7	≤920	

Damages due to tropical cyclones are caused by high wind speeds, storm surges, heavy rain and spawning tornado activity (Zehnder, 2021). Strong sustained winds and wind gusts can cause catastrophic damages to e.g., house, trees, and electricity infrastructure. Already a TC of category 1 can cause damage to roofs, shingles, and gutters while a TC of category 5 will destroy a high percentage of framed houses with total roof failure and wall collapse. However, large amounts of the damage are often due to accompanying storm surges which are also responsible for most of the deaths attributed to landfalling TCs.

The risk associated with natural hazards such as TCs can be understood as an interaction of hazard, vulnerability and exposure (IPCC, 2014). The hazard itself is a combination of the probability of the hazard occurring and its intensity. The analysis of TC landfall rates within this thesis mostly concerns the probability component when assessing TC hazard risk.

The main challenge in correctly assessing the hazard component associated with TC risk is the short historical data record which is not sufficient to calculate risk of extreme events with high return periods. For example, to calculate the risk of a 200-year event a data record of at least 10,000 years would be necessary. Contemporary TC hazard risk analysis mostly relies on catastrophe models that expand the short data records with a suite of thousands stochastic events. Stochastic events are simulated, physically realistic storm tracks that are created by statistically extrapolating historical records of past storms (Philip *et al.*, 2019). While they include some statistically added fluctuations, their characteristics and parameters are largely similar to past observed events. This is due to the fact, that statistical models strongly rely on distribution of past observations while the underlying parameters generating those observations and the associated uncertainties are unknown (Woo, Maynard and Seria, 2017). This can lead to a systematic underestimation of TC probabilities which is also illustrated by ever more records being set within the last few years. Only recently a very active storm season in the North Atlantic with 30 named storms developing in 2020 kept everyone uneasy (Bertogg, 2021).

2.2   Tropical Cyclone Forecast Data as Counterfactuals

Within the last few decades prediction of TC tracks has improved remarkably due to progress in research and development of numerical weather prediction (Yamaguchi, Nakazawa and Hoshino, 2012). Since around 1990, multilevel global and regional dynamical models have become increasingly more accurate and replaced statistical or statistical dynamical models (Rappaport *et al.*, 2009). Keys in this advancement were better assimilation of satellite data, improved model physics and higher model resolution. Another improvement in TC track predictions is due to the implementation of the ensemble prediction system. Instead of making a single forecast a set, or ensemble of forecast is produced to indicate the range of possible future states of the atmosphere. Ensemble TC track predictions perform better on average and are able to capture observed tracks that single deterministic predictions may miss.

For each TC several forecasts containing several ensemble tracks are produced. This is a ready to use data base containing possible alternative track outcomes for past TCs. They are based upon the dynamic situation at several points in the development of the TC and account for slight differences in the atmospheric conditions. Compared to stochastically simulated TC tracks they are less biased towards the outcome of past events. Their range and distribution concerning TC variables are not guided by the same underlying assumptions about the distribution that is based in a short historical record.

At each timestep the spread of the forecast members changes. Early storm forecasts will deviate stronger from the observed track record while forecast from later stages, e.g., closer to landfall, will more likely encompass the observed path of a storm. In this sense the probability of a forecast member representing the actual outcome changes with lead time. This is important when assessing the quality of forecasts. However, in counterfactual risk analysis the probability of occurrence associated with individual ensemble members is of less importance. The focus is rather how the observed TC could have turned out if the conditions were only slightly different. In that sense one can argue that each forecast member, independent on lead time, is a plausible alternative outcome of a past event, or a counterfactual.

3 Methods and Data

To investigate alternative outcomes of past tropical cyclones, existing forecast data is used as source for counterfactual information. Bayesian statistical methods are applied to understand how the landfall distribution based upon forecast data changes from a distribution using only observed track data.

The analysis is conducted for all six major TC basins. For each basin probability distribution of yearly landfall rates are analyzed aggregated across the basin as well as for each impacted country of the region individually. As the EP mainly is of interest for the USA, in this case the different US states are analyzed individually.

3.1 Methods

Statistical inference is concerned with learning something about a population based on data sampled from the population (The CTHAEH, 2016). The main goals are estimation of unknown parameters and data prediction. Bayesian inference is to model a set of data with a distribution depending on unknown parameters. It relies on Bayes' Theorem which relates the probability of a parameter given the available data or observations to the model of the observations. The understanding of probability in Bayesian inference is rooted in terms of physical tendencies and degrees of beliefs rather than just the long-term frequencies on which frequentist inference is based on. Thus, probability can be seen as a measure of belief of a certain event occurring. The shape of the distribution expresses the belief as well as the attributed uncertainties for the belief in possible parameter values.

3.1.1 Bayesian Inference of TC Landfall Rates

This thesis models yearly landfall rates of TCs on observed landfall rates or counterfactual landfall rates based on forecast data. Yearly landfall rates can be represented as simple count data and thus be modelled as a Poisson random variable (Elsner and Bossak, 2001). We assume that observed yearly landfall (y) follows a Poisson mass distribution

$$y \sim Poisson(\lambda) \tag{3.1}$$

The probability of an observation y is

$$p(y|\lambda) = \frac{\lambda^y e^{-\lambda}}{y!}, for\ y = 0,1,2,\dots \tag{3.2}$$

λ is called a parameter, or rate, of the distribution controlling the distribution's shape. It can be any positive real number. y on the other hand, must be a non-negative integer. A useful property of the Poisson mass distribution is that its expected value is equal to its parameter, i.e.:

$$E[y|\lambda] = \lambda \tag{3.3}$$

meaning λ represents the average number of yearly landfalls or mean landfall rates. With λ being the single parameter governing the distribution of yearly landfalls the Bayesian inference analysis is defined by modelling the belief in λ. While the value of λ is unknown, Bayesian inference attributes a probability function to it based upon the knowledge and observations of

yearly landfall rates. The goal is to arrive at a probability distribution representing the probability of all possible values for the unknown parameter λ. Put into Bayes' formula it follows

$$p(\lambda|y) = \frac{p(\lambda)p(y|\lambda)}{p(y)} \tag{3.4}$$

which reads as the probability of the parameter λ given the observations y to the model $p(y|\lambda)$ that we assume for the observations. $p(\lambda)$ is the prior distribution for λ which represents the initial guess or previous knowledge about the parameter. Bayesian inference is to update the initial belief based on the available observations. The resulting probability distribution for λ is called the posterior distribution, $p(\lambda|y)$. The marginal distribution, $p(y)$, is independent of the parameter λ which means it is a constant and can be omitted:

$$p(\lambda|y) \propto p(\lambda)p(y|\lambda) \tag{3.5}$$

Prior

For a model using a Poisson distribution it can be shown analytically that the prior as well as the posterior distribution of the parameter governing the process can be represented in a gamma distribution (e.g. see Donovan and Mickey, 2019; Gelman et al., 2020). In Bayesian terms, the gamma distribution is used as a conjugate prior for the Poisson distribution. The prior distribution represents the belief in each possible value for λ before introducing new observations. Also, it expresses the model of the data to generate the posterior distribution, $p(\lambda|y)$ (Donovan and Mickey, 2019). After deriving the likelihood $p(y|\lambda)$, which is of the form $\lambda^a e^{-b\lambda}$, it follows, that the prior distribution must be in the form

$$p(\lambda) \propto e^{-\beta\lambda}\lambda^{\alpha-1} = Gamma(\alpha, \beta), \tag{3.6}$$

which is a gamma density with hyperparameters α and β (Gelman *et al.*, 2020). The gamma distribution is a continuous probability distribution of the probability density function. The hyperparameters are referred to as shape (α) and the rate (β). They are the guiding parameters of the gamma distribution. As such they not only represent the previous knowledge about λ but they also express the uncertainties regarding the belief in $p(\lambda)$ representing the true value of λ. They are related to λ through the following formula:

$$\lambda = \frac{\alpha}{\beta} \tag{3.7}$$

Posterior

The gamma distribution is a conjugate distribution which can be updated with Poisson distributed data. The resulting posterior distribution is also a gamma distribution with updated hyperparameters. The effect of the new information or data used is then expressed in terms of changes in parameter values (Donovan and Mickey, 2019). It follows the posterior distribution

$$p(\lambda|y) \sim Gamma(\alpha + n\bar{y}_i, \beta + n) \tag{3.8}$$

where $y_i = (y_1, \ldots, y_n)$ is a vector of independent and identically distributed observations. n is the number of years observed. By modelling TC landfall rates in this manner, it is possible to arrive at a posterior distribution mathematically.

Posterior Predictive (PP) Distribution

Based upon the posterior belief for λ, the next step is to predict future landfall rates, $\tilde{y}$. While observed landfall rates follow the distribution $p(y|\lambda)$, the true value for λ is not known. To account for the uncertainties, we average over all possible values of λ to get a better idea of the distribution of yearly landfall rates. The basic idea is to randomly sample possible values for λ from the gamma posterior distribution and then sample random values for yearly landfall rates from a Poisson distribution based on the sampled value for λ. Analytically it can be shown that the resulting posterior predictive distribution follows a negative binomial distribution with the following parameters (Gill, 2015):

$$\tilde{y} \sim NB\left(\alpha + n\bar{y}, \frac{1}{n + \beta + 1}\right) \tag{3.9}$$

3.1.2 Practical Calculations

Within this analysis the prior distribution is built by fitting a gamma distribution to observed yearly landfall rates up to 2007. The observed mean landfall rates for this period, μ_{hist}, thus represent the prior estimate for λ:

$$\mu_{hist} = \frac{y_{2007-n} + y_{2007-(n+1)}, \ldots + y_{2007}}{n} \tag{3.10}$$

By choosing values for the hyperparameters shape (α) and the rate (β) it is expressed how high the certainty is that μ_{hist} represents the true value of λ. The hyperparameters are chosen in a way that the relationship expressed in equation (3.7) is met:

$$E[\lambda] = \mu_{hist} = \frac{\alpha}{\beta}. \tag{3.11}$$

Figure 3.1 illustrates how different values for the hyperparameters change the certainty for λ. Narrow distributions imply a very high certainty for our belief in λ, while a broad distribution implies high uncertainties for the mean landfall rates based solely on data up to 2007. Within this analysis several values for scale and rate were tested but finally the parametrization of the prior distribution was chosen as follows:

$$\alpha = \mu_{hist} * \beta \text{ and } \beta = 4.$$

As the dark red curve in Figure 3.1 shows, these values express some confidence in previous landfall rates while still expressing a rather high level of uncertainty acknowledging the fact that the data used for the prior landfall rates is very short. Also, the data is influenced by available data quality which differs among regions, time periods and providing agency. For the analysis across basins, countries, and providers to be comparable the same values for α and β were used for all the analysis represented. It should be noted however, that is a strong simplification, as one cannot assume that the uncertainty is the same for all regions, countries

14

or time periods used for the calculation of μ_{hist}. However, this allows for a higher comparability across case studies.
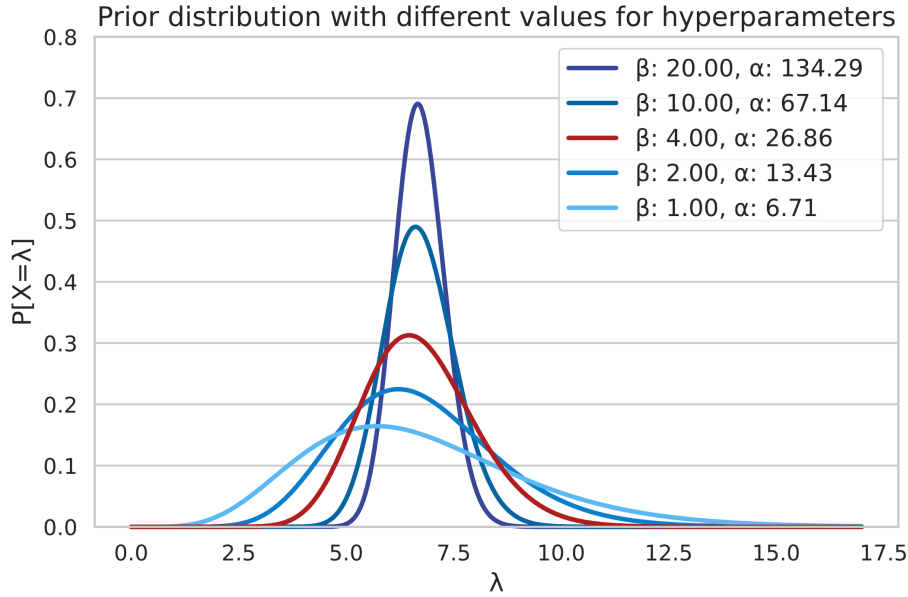

Figure 3.1: Gamma probability mass distribution depending on chosen hyperparameters. Illustrated based upon μ_{hist} for the NA using data from 1980 to 2007. For all analysis presented in this thesis the hyperparameters used are $\alpha=\mu_{hist}*\beta$ and $\beta=4$ (dark red line).

In a next step, the prior distribution is updated by incorporating new information of landfall rates from 2008 to 2019. On one hand this is done using observed landfall rates as a reference distribution. On the other hand, the posterior belief is updated with counterfactual landfall rates based on forecast tracks. A scenario of counterfactual landfall rates represents one in many possible combinations of all available forecast members. With the large number of forecast tracks for each TC in the period 2008 to 2019 there can be as many posteriors built as there are forecast track combinations.

The mean of the gamma posterior distribution represents the expected value for λ based upon the added information, $E[\lambda|y] = \frac{\alpha+n\bar{y}_l}{\beta+n}$. The variance of the distribution is $\frac{\alpha+n\bar{y}_l}{(\beta+n)^2}$. The posterior distribution is denoted as $p(\lambda|y_{obs})$ for the posterior belief based on observed landfall rates for the period 2008-2019. The posterior belief based on the worst-case scenario with the highest possible numbers of landfalls is denoted as $p(\lambda|y_{CFmax})$. The respective mean values of the posterior distributions are $E[\lambda|y_{obs}]$ and $E[\lambda|y_{CFmax}]$.

The posterior predictive distribution based upon the observed or counterfactual landfall rates portrays how future landfall rates, $\tilde{y}$, would be expected given the respective scenario. Their distributions for observed and worst-case scenarios are denoted as $p(\tilde{y}|y_{obs})$ and $p(\tilde{y}|y_{CFmax})$ respectively.

The negative binomial posterior predictive distribution has the same mean value as the posterior distribution, $\frac{\alpha+n\bar{y}_l}{\beta+n}$. The variance however is greater for the posterior predictive distribution and

defined as $\frac{\alpha + n\bar{y}_t}{(\beta + n)^2}(\beta + n + 1)$. This is due to the additional uncertainty based on the fact, that we are sampling new data values.

3.2 Data

For the Bayesian inference analysis of counterfactual information two datasets are necessary. A dataset of observed TC tracks is necessary to build a prior knowledge or belief of past landfall rates as well as the reference for the counterfactual scenarios. A dataset of TC forecast tracks is used as a basis to build different counterfactual scenarios.

3.2.1 Observed Tropical Cyclone Data

The most complete set of tropical cyclone data available is the International Best Track Archive for Climate Stewardship (IBTrACS) dataset from the National Oceanic and Atmospheric Administration (NOAA) (Knapp *et al.*, 2010). It contains a global collection of best-track data combining data from different agencies worldwide. Best-track data of a tropical cyclone contains the best estimate of storm position and intensity at intervals of 6 hours. If there is data available from several agencies for one single event, the information is combined using objective techniques that account for the differences between the international agencies (Knapp *et al.*, 2010).

As historic TC data this thesis uses data from the IBTrACS Project, Version 4 for the NA, EP, WP, NI, SI and SP basin (Knapp *et al.*, 2018). As the IBTrACS dataset is a collection of best-track data rather than a reanalysis the data available is strongly influenced by methods used by different agencies and time periods. For some regions the dataset contains TC data back to 1848 but global records spanning all basins only date back to 1945. TCs originally were mostly of interest for shipping and only in the late 1950s and 1960s there was an increased interest in their climatology and the risk they pose for coastal communities (Knapp, 2019). In the same period routine aircraft observations were introduced, improving location estimates. Another big improvement in TC reporting came with the incorporation of data provided by satellites. First meteorological satellite observations in 1960s were merely able to identify systems from space. However, with routine microwave imager satellites starting in the 1980s observations on rain structure, expanse of winds, as well as the eye position improved considerably. Thus, TC data from 1980 on is considered to be the modern era, as geostationary satellite coverage was nearly global and global coverage from polar orbiting data was more widely available.

To account for this difference in data availability and quality and to gain more insight into the sensitivity of the analysis on the prior information, two time periods are compared: 1950-2007 and 1980-2007. All BI analysis is carried out using prior information based upon the two periods individually.

One problem when counting storms in IBTrACS is that the operational procedures when including storms in TC reporting are dependent on the different agencies and may have changed over time. For example, some agencies may include tropical depressions and sub-tropical storms while others do not. To account for this fact, within this analysis only named tropical cyclones are included. Not named storms are in general cyclones that never reach an intensity higher than a tropical depression and therefore usually are not analyzed in much detail. Thus,

by only including named storms some of the mentioned issues can be avoided. However, for future analysis more detailed analysis of the included TCs might need to be considered (see e.g. Schreck et al., 2014). Another, more practical reason for only including named storms is also the problem of matching not named TCs across the different datasets of historic and forecast data.

The IBTrACS dataset for the NA basin includes 1078 tracks within the period 1950 – 2019. 748 of those tracks are named tropical cyclones. For the other basins the numbers are similar (see Table 3.1). An exception is the NI basin. For this basin named tracks only exist from 2005 onward. For this reason, the analysis of the NI basin includes all available tracks.

Table 3.1: Number of IBTrACS tracks and landfalls per basin and period.

		NA	EP	WP	NI*	SI	SP
IBTrACS dataset 1950-2019	all tracks	1078	1228	2040	608	1166	798
	named tracks	748	971	1728	55	845	514
	landfalling tracks	462	334	1421	479	402	380
		61.76%	34.40%	82.23%	78.78%	47.57%	73.93%
Prior period 1 1950-2007	named tracks	571	749	1440	543	691	413
	landfalling tracks	360	245	1171	428	336	311
		63.05%	32.71%	81.32%	78.82%	48.63%	75.30%
Prior period 2 1980-2007	named tracks	318	484	701	137	549	366
	landfalling tracks	188	148	583	114	186	195
		59.12%	30.58%	83.17%	83.21%	33.88%	53.28%
observation period 2008-2019	named tracks	177	222	288	65	154	101
	landfalling tracks	102	89	250	51	66	69
		57.63%	40.09%	86.81%	78.46%	42.86%	68.32%

* As NI basin only has named tracks starting in 2005 all tracks are included for this basin.

For the analysis of the variable landfall only the positional information of latitude and longitude is needed. Landfall occurs when the area of strongest wind i.e., the eyewall, crosses the coastline or comes within 50 km of it. The variable used in the Bayesian Inference are yearly landfall rates. Those are simple counts of tracks making landfall within each year and are calculated for the entire basin as well for all relevant countries individually. Figure 3.2 gives an overview over the yearly landfall rates of the entire dataset across the different basins.

As for the SI and SP the main TC season is in winter the landfall rates are not calculated yearly but rather for the TC season. For both basins a TC season spans TC tracks from August of the previous year until July of the indicated season year. Thus e.g., when mentioning the reference period from 2008-2019 for the SI and SP basins this in fact refers to seasons 2009-2020 including data from August 2008 until July 2020.
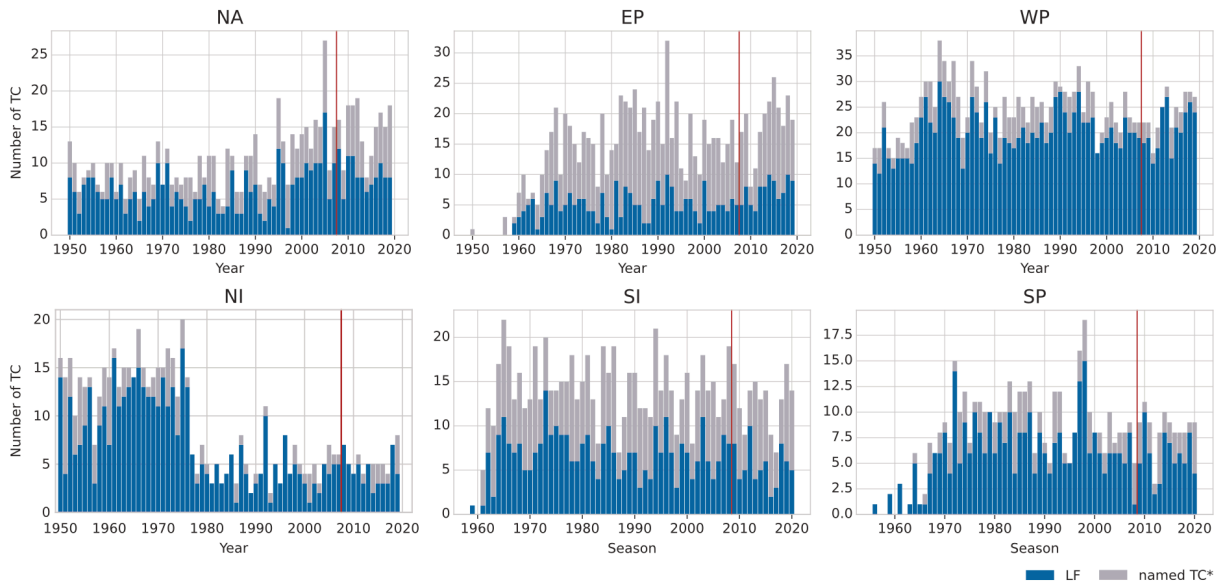
Figure 3.2: Bar chart of yearly number of named TCs (grey) and landfall rates (blue) for the individual basins. Red line indicates the cut-off year or season for the data used to calculate the prior distribution (left of red line). Data on the right of the red right are used to calculate the reference posterior distribution (observation period). *NI dataset also not named tracks are included.

3.2.2 Forecasting Data and Counterfactual Scenarios

The forecast data used within this thesis is part of The Observing System Research and Predictability Experiment (THORPEX) which is a big component of the World Weather Research Programme under the World Meteorological Organization. The THORPEX Interactive Grand Global Ensemble (TIGGE) was initiated in 2005 and contains among many more forecasting data a dataset of tropical cyclone track data (Bougeault *et al.*, 2010). This dataset called TIGGE Model Tropical Cyclone Track Data (further simply referred to as TIGGE dataset) holds model analysis and forecast data from several international meteorological agencies. The dataset contains track data since 2008 and is updated daily (National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce *et al.*, 2008).

For each TC event and provider there are several forecasts available, which in turn contain several forecast members. Also, the number of available forecasts and members varies greatly among TC events. The available forecast track variables depend upon the providing agency but in general include name, time, position (latitude and longitude), minimum pressure, maximum wind at 6-hourly intervals after the start of the forecast. Although there are TC forecast data available from several agencies world-wide, only the data from two agencies span the entire period from 2008-2019 without considerable data gaps. Thus, only data from the following providers are used and compared in the data analysis:

- **KWBC** - National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Comerce and Meteorological Service of Canada, Environment Canada
- **ECMWF** – European Centre for Medium-Range Weather Forecasts

To be able to compare the different providers for each basin a set of counterfactual histories is created based upon forecast members from the KWBC, the ECMWF or the combined dataset using tracks from both providers. Table 3.2 gives an overview over the available number of events and tracks within the TIGGE dataset. For example, the KWBC provides for the NA basins 171,684 forecast tracks that span 432 storm events for the period between 2008 and 2019. To be able to create credible counterfactual landfall rates for the period 2008-2019, the track data used must be comparable to the dataset of observed landfall rates in the same period. Thus, only forecast tracks for named storms in the IBTrACS dataset were selected. Next to named TC events the TIGGE dataset also contains forecasts for not named TCs as well for invest tracks. Invest tracks are forecast for areas of disturbed weather, where possibly a TC could develop and thus are monitored. After the selection according to named observed tracks (IBTrACS), the number of NA tracks is reduced from 84,061 tracks (Table 3.2). To arrive at a robust dataset several checks and corrections had to be implemented as described in the following.

Table 3.2: TIGGE - Number of events and forecast tracks. The selected tracks show the sum of all forecast members across the selected named TC events. The events are selected by name and cross-referenced with observed TC of the same period. The difference of the selected and observed events indicates the number of missing events in the TIGGE dataset.

Basin	Provider	Number of TIGGE events	Number individual tracks	Selected (observed) events	Selected tracks
NA	ECMWF	314	110596	176 (177)	85337
	KWBC	432	171684	156 (177)	84061
EP	ECMWF	417	126291	222 (222)	104472
	KWBC	516	185061	182 (222)	96258
WP	ECMWF	333	169134	281 (288)	149901
	KWBC	681	237864	254 (288)	131631
NI	ECMWF	64	17218	48 (65)	14162
	KWBC	104	23105	33 (65)	14899
SI	ECMWF	374	128051*	153 (154)	62553
	KWBC	375	120251	144 (154)	68391
SP	ECMWF	374	128051*	100 (101)	40657
	KWBC	229	75403	94 (101)	47506

* Within the data provided by the ECMWF tracks for the SI and SP basins are compiled in a single dataset.

Treatment of TIGGE Data and Data Problems

The TIGGE database is a collection of forecast ensembles from different agencies. As such, they make available data on TC forecast tracks as provided by the agencies and are not further homogenized or otherwise quality controlled. This means that differences across basins, time periods and providers are not accounted for. As within this thesis only the track or path is relevant only the variables time and location are included. Systematic errors in track geometry are difficult to detect and correct for. The biggest obstacle is, that a visual check of the track information is unavoidable. However, looking at the sheer number of tracks which is overall several hundred thousand tracks this cannot be done for each single track. Thus, there is no thorough analysis or homogenization implemented in the presented analysis. However, some

simple quality controls are conducted. Tracks meeting the following criteria are excluded from the data analysis:

- duplicate tracks
- tracks with 2 or less timesteps
- tracks located out of bounds of basin
- tracks with unnatural changes in location between 2 timesteps (chance > 10°)

The tracks are grouped by event name and time to create individual events. For each named event in the TIGGE dataset a visual check allows for correction of the following issues:

- reversed latitude
- wrong attribution of event name

Finally, the grouped dataset is cross referenced with the IBTrACS dataset to select all observed, named, tracks within the period. The problem with this selection by name, is that possible forecasts from an earlier stage of an event will not be included, as they are not yet named correctly. For example, if an invest tracks later develops into a named storm not all forecasts from earlier stages will be attributed with the later storm name. However, due to the large number of events in the TIGGE dataset a correction of all not named or invest is not conducted.

Further, for most basins not for all observed events forecast tracks is available (see Table 3.2). For some regions and periods there are gaps in the available data, especially in the earlier years of the period. For example, for the years 2008 and 2009 there are several months of data missing from the KWBC dataset. In the NA for example, out of the 177 IBTrACS events only for 156 events are forecast tracks available in the KWBC dataset. A much bigger gap exists for the NI basin. As for this basin, also non-named tracks had to be included there are several events where no forecasts are available.

To still be able to achieve a full set of events, the missing events in the TIGGE dataset were replaced with the observed IBTrACS track. This means however, that for those events no counterfactual tracks are available and only one outcome (landfall or no landfall) could be considered.

Creating Counterfactual Scenarios

Landfall is calculated for the whole basins as well as for individual countries for each forecast member. Counterfactual scenarios of landfall rates are created by sampling a forecast track for each observed track in the period 2008 – 2019. The resulting yearly landfall rate represents one possible counterfactual scenario of how the landfall could have been distributed within the observed period. As with random sampling of the forecast tracks not the entire range of possible counterfactual scenarios could be covered the counterfactual scenarios were created manually.

To this purpose, for each subset of individual provider and basin or country the probability of an event making landfall is calculated based on the number of landfalling tracks compared to the total number of forecast tracks:

$$p(landfall) = \frac{number\ landfalling\ forecast\ tracks}{total\ forecast\ tracks}$$

With the thesis' focus on downward counterfactuals the scenario of most interest is the worst-case scenario. This scenario assumes landfall for an event if at least one of the forecast tracks make landfall. Thus, all events with landfall probability higher than 0 were included. To then build a range of possible counterfactual scenarios the events were sorted by their landfall probability. For each new history the event with the lowest probability of landfall is assumed to not make landfall. The final scenario represents the best-case scenario and assumes only landfall for events where the probability of landfall is equal to 1, meaning all forecast tracks make landfall.

This thesis focuses on exploring counterfactuals in terms of their range and focuses on what extremes downward counterfactuals can reveal. Therefore, a further analysis of the probabilities regarding a TC making landfall based on the percentage of landfalling forecast tracks is not included in this thesis.

4 Results and Discussion

As described above, the Bayesian inference of yearly landfall rates are conducted for the different …

… prior distributions based upon the period (1950-2007 vs 1980-2007),

… providing agencies of counterfactual data (forecast data from KWBC vs. ECMWF vs. KWBC+ECMWF)

… basins and countries (NA, EP, WP, NI, SI, SP)

The focus of the results presented, is how the incorporation of counterfactual TC data change our belief for yearly landfall rates in comparison to the same Bayesian Inference using observed TC data.

For the individual basins and countries not all results across case studies are shown. For the analysis across basins, an extended version, including results of each case study across prior period and provider can be found in Appendix A.

4.1 Observed TC Landfall Rates

As seen in Figure 3.2, the data availability is not the same for all basins. For the NA and the WP TC best-track data are available all the way back to 1950. For the SI, SP and EP however, data is only reliably available since the 1960s. For the NI best-track data goes back to the 1950s, but the dataset contains an abrupt shift in the number of TC events around 1980. This shift is strongly informing the prior distribution, as the mean values for the two compared periods differ by 3.3 landfalls a year (Table 4.1). However, for none of the other basins the differences in mean values are bigger than 1. With such small differences the period used to form the prior distribution does not have a great impact on the BI analysis. This is tested by updating the prior probability by adding new observed landfall rates for the period 2008 to 2019. The right side

of Table 4.1 shows how adding observed landfall rates influences the posterior belief in λ. The comparison of posterior distributions depending on different priors shows very little differences across the two case studies. In fact, they are even smaller than the prior mean values, as by introducing new information the distribution shifts towards the newly added data (Figure 4.1).

Table 4.1:Comparison of prior mean values of yearly landfall rates based on periods and difference between them (left). Comparison of resulting posterior belief in λ, based on updated belief using observed TC landfall rates 2008-2019 as well as difference of the two posteriors based on different priors (right).

Basin	Prior			Posterior observed		
	μ_{hist50} 1950-2007	μ_{hist80} 1980-2007	$\Delta\mu_{hist}$	$E[\lambda\|y_{obs}]$ 1950-2007	$E[\lambda\|y_{obs}]$ 1980-2007	$\Delta E[\lambda\|\lambda_{obs}]$
NA	6.207	6.714	0.507	7.927	8.054	0.127
EP	4.824	5.286	0.462	6.768	6.884	0.116
WP	20.190	20.821	0.632	20.672	20.830	0.158
NI	7.379	4.071	-3.308	5.032	4.205	-0.827
SI	6.857	6.643	-0.214	5.839	5.786	-0.054
SP	6.347	6.964	0.617	5.899	6.054	0.154

Again, the exception is the NI basin, which is strongly influenced by the big shift in observed landfall rates. Noticeable across the basins is that except for the NI and SI basins, the mean value in the prior period starting 1950 is lower than for the period starting 1980. Comparing the posterior belief across basins shows further, that by adding data from 2008 – 2019 the probability distribution shifts towards higher mean yearly landfall rates for all basins except the SI and SP basin. For the NI basin the same trend is only observed for BI based on prior data incorporating data from the modern era (μ_{hist80}). This suggests, that for the northern hemisphere there seems to be an increase in yearly landfalls across the compared periods. The same patterns can be observed looking at the prior distributions for the individual countries. With a few exceptions such as India or Bangladesh most mean values are higher for the $p(\lambda_{80})$ and mostly the differences are relatively small. The main exception is India with a difference $\mu_{hist80} - \mu_{hist50} = -3.22$. All other countries show differences of less than 1 landfall a year. Also, the trends in the posterior beliefs show a tendency towards higher beliefs in λ for the northern hemisphere. Although there are a few exceptions, none of them show a big decrease in yearly landfalls. When updating the belief using counterfactual data, for the worst-case scenario the posterior distribution, $p(\lambda|y_{CFmax})$, shows the same patterns as $p(\lambda|y_{obs})$. Also, here the differences are relatively small for all basins (Table A 1).

As mentioned in chapter 3.2.1 the data availability for the different basins largely depends upon reporting practices of the different agencies contributing to the IBTrACS dataset (Knapp, 2019). Most prominent is big shift in yearly TC numbers for the NI basin. This can be explained by the selection upon which TCs are included in the IBTrACS. Up to around 1979 all TCs, tropical storms and tropical depressions are included in the dataset. From around 1980 onward only TC or tropical storms are included in the dataset, while tropical depressions are no longer included. This problem however only occurs for the NI ocean as usually tropical depressions

are not named and thus not included. But as mentioned for the NI all TCs had to be included, as they were not named up until 2005.
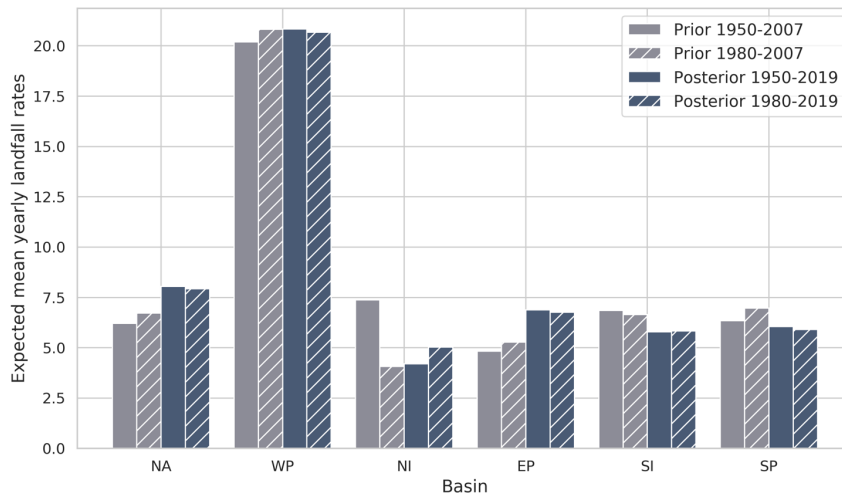

Figure 4.1: Comparison of average yearly landfall rates for two prior periods 1950-2007 and 1980-2007. Grey: mean landfall rate of the observation period. Blue: expected mean landfall for the posterior distribution based on observed landfall rates 2008-2019.

For all basins it needs to be considered, that not all regions had the same amount and quality of TC observations especially before the modern era of widely available satellite data starting around 1980. This means on a global scale the data from the modern era are more reliable, especially when comparing different basins. This difference in data availability could also explain the fact, that the average landfall rate is higher for the modern observation era. Before satellite observations there certainly were more missed storms. Thus, just by comparing those two periods one cannot make any assumptions about trends in number of TCs.

When looking at the posterior distribution based on observed IBTrACS data (Table 4.1), it becomes noticeable, that for most basins the posterior belief shows a shift towards higher expected values for λ. One might argue that TC forecasts are constantly improving and the period after 2008 still has a lower number of missed storm, than the period from 1980-2007. Another explanation could be that there was a change in the practice of naming storms. However, this could be different for different basins and providers and therefore difficult to reconstruct. When assuming that there are no significant changes in practices and quality of the best-track data, these results suggest an increase in landfalling TCs for the basins in the northern hemisphere. The IPCC Assessment Report 6 attributes low confidence towards most reported long-term trends due to the heterogeneous character of the best-track data (IPCC, 2021). As for the global frequency of TCs the IPCC suggest that it will likely remain essentially unchanged or decrease. This is however on a global scale and does not directly apply for the landfall frequency as well. The landfall frequency might also be dependent on shifts in the tracks itself. For example, Lackmann (2015) illustrated, that in future projections for Hurricane Sandy the landfall would be expected more than 200 km further north-northeast than the observed landfall in New Jersey.

So, whether there is a clear trend or not, this is where the method of BI can be of high value as it incorporates the notion of learning from new data and incorporating multiple hypothesis for λ by expressing our belief in a probability distribution.

Overall, for most basins the differences in the prior distribution based on the two compared

periods are relatively small. Considering the knowledge about quality and completeness of the data it must be assumed that data from the modern era is of better consistency than pre 1980 data. Thus, in the following only results from BI analysis using prior information from 1980-2007 are considered. As such, the regime shift in the NI basin does not influence the results and for all basins the data availability on which the prior is bases is comparable.

4.2 Counterfactual Scenarios

As described in chapter 3.2.2 downward counterfactual histories are built by incorporating all possible landfalls for the worst-case scenarios. The best-case scenarios however only incorporate landfalls for TCs when all forecast members make landfall. Table 4.2 gives an overview over the historic mean landfall rates and the sum of landfalls for the observed period.

Table 4.2: Comparison of average landfall rates and number of landfalls for the observation period 2008-2019 for range of counterfactual scenarios based upon data from different providers. The best-case scenario is built using a combination of forecast members of all events arriving at the lowest mean landfall rate. The worst-case scenario is the combination leading to the highest mean landfall rate.

	IBTrACS			KWBC		ECMWF		KWBC+ECMWF	
Basin	mean LF	sum	CF scenario	mean LF	sum	mean LF	sum	mean LF	sum
NA	8.50	102	best-case	2.75	33	2.58	31	2.25	27
			worst-case	11.00	132	10.83	130	11.67	140
EP	7.42	89	best-case	1.50	18	1.00	12	0.83	10
			worst-case	14.75	177	14.58	175	15.58	187
WP	20.83	250	best-case	7.08	85	5.83	70	4.33	52
			worst-case	23.42	281	23.42	281	23.58	283
NI	4.25	51	best-case	2.75	33	2.92	35	2.17	26
			worst-case	4.58	55	4.83	58	5.00	60
SI	5.50	66	best-case	1.33	16	1.92	23	1.08	13
			worst-case	11.83	142	10.58	127	12.50	150
SP	5.50	69	best-case	1.33	16	1.42	17	1.17	14
			worst-case	8.00	96	7.67	92	8.25	99

It compares the range from the worst- to best-case scenarios for the different providers KWBC, ECMWF, and the two combined. Combining the two datasets achieves the largest range of counterfactual scenarios. Due to the higher number of forecast tracks considered it is more likely that at least one forecast member of an event will make landfall. As such, when looking into downward counterfactuals when combining the two datasets more and higher extremes can be produced. However, the differences across the three case studies are in general relatively small consisting of a few events more or less making landfall. Figure 4.2 illustrates landfall sums across the period for the worst-case scenario across different case studies. For example, the worst-case scenario for the SP basin includes a sum of 96 landfalls for the KWBC dataset and 92 landfalls for ECMWF. When combining the two dataset a total sum of 99 landfalls can be achieved. The biggest difference is seen for the SI basin where the KWBC and ECMWF count 142 and 127 landfalls respectively and the combined dataset 150. Overall, the influence

of the data from different providers on the counterfactual analysis is relatively small (Appendix A).
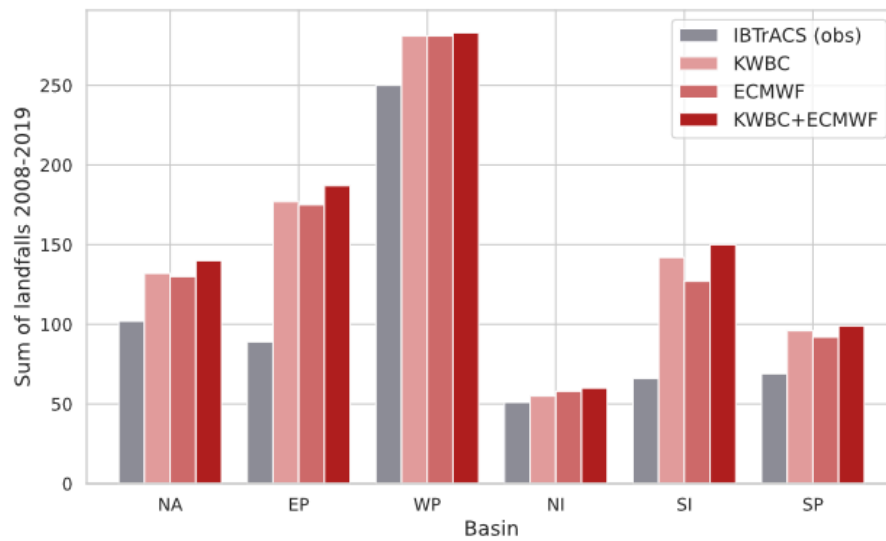

Figure 4.2: Comparison of worst-case scenarios across providers. Sum of landfalls for the period 2008-2019. Observed landfalls as a comparison in grey.

The range of the counterfactual scenarios from best to worst-case is very large and shows large deviations towards the observed landfall rates. While there are many better case scenarios where some of the landfalling TCs could have potentially missed the coast, there are also many TCs where some or at least one forecast member predicted a possible landfall. The sum of all landfalls as well as the mean value for the worst-case scenarios across the different providers show a large increase compared to the observed landfall rates. Compared to the reported 89 landfalls in the EP basin between 2008 and 2019 (mean of 7.42 landfalls per year), the counterfactual worst-case scenario of the combined dataset with its 187 landfalls (mean of 15.58) shows more than a doubling of the yearly landfall rates. Lower values are observed for the NA, WP, and NI basins (Table 4.2). For example, in the NI basin there were 51 landfalling TCs observed, while the worst-case scenario counts 11 more landfalling TCs.

The small differences across providers reflect the fact, that both the KWBC and ECMWF are well established forecast centers producing comparable forecasts data for all TC basins. A more thorough analysis on an event level however might reveal more differences in the forecasting and possible landfall of individual events. Such differences are to be expected as the two forecast agencies have their own models with different parametrizations. One such example illustrating these differences was Hurricane Sandy in 2012. Sandy took an unexpected left turn making landfall in southern New Jersey leading to large casualties and damages. This left turn was not detected in the early forecasts from the US agencies, while the ECMWF model did show earlier forecast members modelling similar trajectories. This difference can be attributed to choice of cumulus parametrization which is different in the European model (Bassill, 2014). When looking into downward counterfactuals and worst-case scenarios it makes sense to include as much possible forecast tracks from different models and providers, as the span of alternative outcome will become broader. This is shown by the larger sums and mean values for the worst-case scenario of the combined dataset. Also, the spatial density of the data can be further increased due to the larger number of forecast tracks and gaps in data availability can be reduced.

Resulting from this vast availability of alternative tracks is also a larger difference between the worst-case and observed landfall rates. The analysis reveals many near-miss events where a TC might have made landfall if the synoptic situation were only slightly different. The accumulation of several worse-case scenarios for those TC could potentially have led to much more landfalls in the past. Depending on the region of the landfalls the resulting hazard damages and losses could be devastatingly worse than what was observed. However, while this thesis mostly focuses on downward counterfactuals one should notice that not all TCs with landfalling forecast members had the same likelihood of actually making landfall. For future analysis, one could look further into the individual probability of each TC making landfall. This would for example mean to incorporate the percentage of forecast members making landfall.

While the worst-case scenarios for the individual basins objectively might not be very probable, they nevertheless are very interesting when exploring the realm of possible future landfalls. While there are no clear trends in the overall frequency of TCs there is substantial literature suggesting trends towards more intense landfalls and an increase in TCs of category 3 and higher (IPCC, 2021). Thus, a downward counterfactual analysis of TC variables can help incorporating extreme storm seasons in risk analysis.

A thorough comparison of the different case studies depending on the providing agency was not done for the individual countries. However, preliminary results show similar features while especially for small states the difference between the two agencies can be larger. This further reinforces the argument towards a combined dataset using as many possible forecast members as available. In this sense, the following analysis of the posterior belief based on counterfactual information and the resulting expected values for future TC landfall rates are presented only for the case study based on the combined dataset from the KWBC and ECMWF.

4.3   Counterfactual Posterior and Posterior Predictive Distribution

In the following the results of the Bayesian inference analysis of counterfactual landfall rates is presented in more detail for all basins. The presented results only include BI using a prior distribution based on landfall rates from 1980-2007 and observations or counterfactuals from 2008-2019. The counterfactual scenarios discussed are based on the combined dataset from the two providers KWBC and ECMWF. Table 4.3 and Figure 4.3 give an overview for the posterior distributions while Table 4.4 and Figure 4.4 illustrate the results for the posterior predictive distribution. The results for the individual countries are summarized in Appendix B. Further, 0 shows the predictive distributions for all countries.
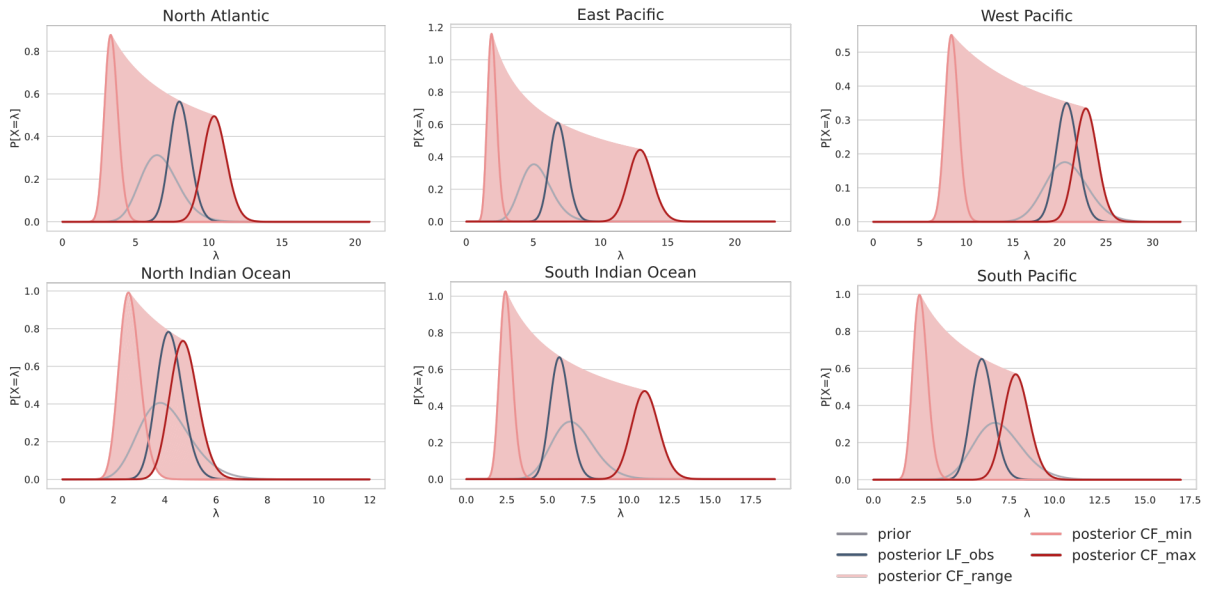
Figure 4.3: Posterior distributions based upon different landfall scenarios for 2008-2019 for all basins. Blue line is posterior distribution based on observed landfall rates. In different shades of red are the individual posteriors based on counterfactual landfall rates from the best-case scenario (left, light red) to the worst-case scenario (right, dark red). Grey line is prior distribution as a reference. All distributions are based on the prior using historic data from 1980-2019 and forecast data from the combined dataset KWBC+ECMWF.
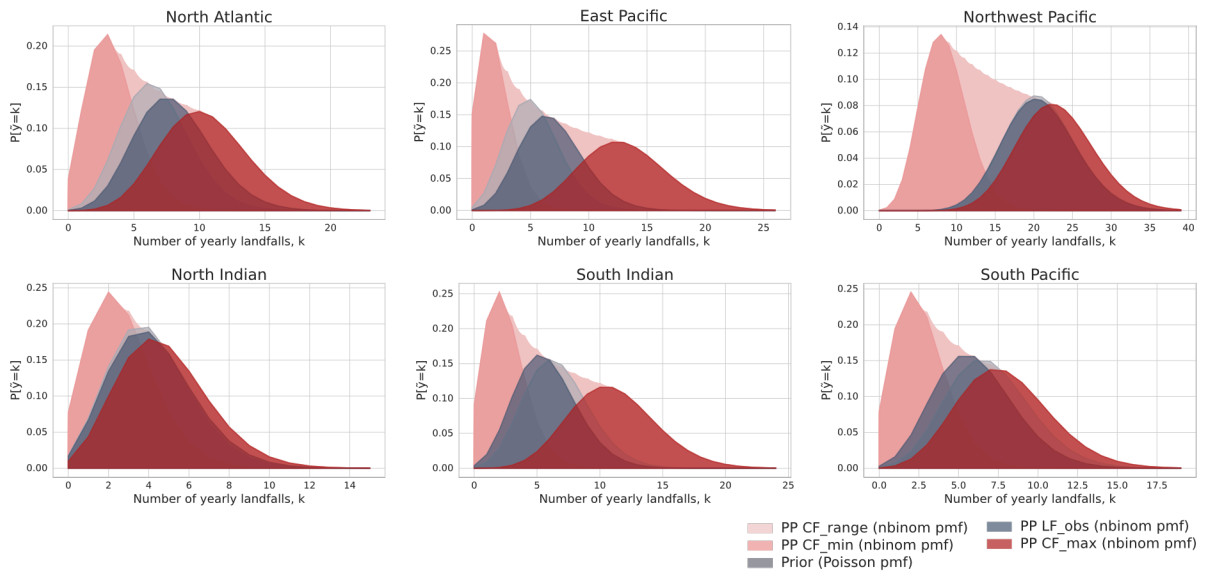
Figure 4.4: Posterior predictive distributions based upon different landfall scenarios for 2008-2019. In blue is posterior predictive distribution based on observed data. In different shades of red are the individual posteriors based on counterfactual scenarios from the best-case scenario (left, light red) to the worst-case scenario (right, dark red). Grey shading is prior distribution as a reference. All distributions are based on the prior using historic data from 1980-2019 and forecast data from the combined dataset KWBC+ECMWF.

Table 4.3: Mean landfall rate for prior period and expected mean landfall rate for posterior distribution from observed data and counterfactual data.

| Basin	Mean prior λ_{hist80}	Mean posterior observed $E[\lambda|y_{obs}]$	Mean posterior worst-case $E[\lambda|y_{CFmax}]$	Difference Posteriors $E[\lambda|y_{CFmax}] - E[\lambda|y_{obs}]$ (Change in %)
NA	6.71	8.05	10.43	2.37 (+29%)
EP	5.29	6.88	13.01	6.12 (+89%)
WP	20.82	20.83	22.89	2.06 (+10%)
NI	4.07	4.21	4.77	0.56 (+13%)
SI	6.64	5.79	11.04	5.25 (+91%)
SP	6.96	6.05	7.93	1.88 (+31%)

Table 4.4: Posterior predictive distribution. Probability of maximum observed event or higher for different posterior predictive distributions (observed vs. counterfactual) and difference (left). Comparison of extreme events at the 95th percentile for the different probability posterior distributions (right).

Basin	Max. observed y 80-07 ($ymax$)	$p(\tilde{y} \geq ymax\|y_{obs})$	$p(\tilde{y} \geq ymax\|y_{CFmax})$	$\Delta p(\tilde{y} \geq ymax)$ [%]	$E[\tilde{y}_{95p}|y_{obs}]$	$E[\tilde{y}_{95p}|y_{CFmax}]$
NA	17	0.002	0.024	2.17%	13	16
EP	10	0.097	0.741	64.44%	12	19
WP	28	0.058	0.129	7.19%	29	31
NI	10	0.005	0.012	0.69%	8	9
SI	11	0.019	0.425	40.61%	10	17
SP	15	0.001	0.010	0.87%	11	13

4.3.1 Results for Individual Basins

North Atlantic

For the NA the posterior distribution based upon observed TCs between 2008 and 2019 already considerably shifts the belief of mean landfall rates towards higher values compared to the prior distribution. This is consistent with Figure 3.2, which shows overall higher numbers of yearly TC events within the last two decades. Out of the 177 TCs for the observation period 102 made landfall (Table 3.1). Considering counterfactual scenarios from different combinations of different forecast members the number of landfalls could have spanned between 27 for the best-case scenario and 140 for worst-case scenario (Table 4.2). Looking at the worst-case scenario, the posterior distribution, $p(\lambda|y_{CF\max})$, shifts significantly towards higher values. The mean of the posterior distribution, or the expected value for λ increases by 2.37 compared to the posterior belief based on observed values ($E[\lambda|y_{CF\max}] - E[\lambda|y_{obs}]$) (Table 4.3). When looking at the distribution for expected future landfall rates, the posterior predictive distribution, $p(\tilde{y}|y_{CFmax})$, we see a considerable increase in probability for higher landfall rates. A good illustration is the year 2005, marking the previous maximum in observed landfalls

($ymax$) with a total of 17 TCs intersecting the US coast. This is an outlier event due to very active storm season. In the PP distribution based upon observed LF rates, the probability of 17 or more landfalls in a year, $p(\tilde{y} \geq ymax|y_{obs})$, is 0.0024. When looking at the worst-case scenario this PP probability increase by a factor of ten to 0.024. Based on the downward counterfactual worst-case scenarios extreme landfall rates in the future would be noticeably higher. An extreme landfall rate at the 95th percentile of the PP distribution, would be expected to be around 13 landfalls a year based on only observed data. For the PP distribution based on the worst-case scenario, such an extreme event would constitute of 16 landfalls a year.

A spatially more detailed analysis of the basin shows the same positive trends in the posterior and posterior predictive distributions for the individual countries (Figure 4.5,Table A 2). The shift towards higher values for the worst-case scenario is more pronounced for the individual countries compared to the basin. Future landfalls rates at the 95th percentile are up to 7 landfalls a year higher for the worst-case scenario. Also, the expected value for λ shows a clear trend for all countries. Absolute numbers suggest a bigger difference in belief for the mean expected landfall rates for larger countries. However, percentage wise the shift is biggest for some of the small island states of the Caribbean, like the Dominican Republic or Guadeloupe where previous landfall data is very sparse. Also, when looking at the increase in the PP probability for the previously highest observed values, some of the smaller states show very big differences.

Figure 4.5: North Atlantic - Difference in expected future landfall rates at 95th percentile of posterior predictive distribution, $E[\tilde{y}_{95p}|y_{CFmax}] - E[\tilde{y}_{95p}|y_{obs}]$. Top right: Zoom to Caribbean.

East Pacific

For the EP only few TC make landfall (Figure 3.2) and the landfall rates are relatively low for this basin. The average yearly landfall up to 2007, μ_{hist}, is 5.29 and the mean value expected based on the reference posterior, $E[\lambda|y_{obs}]$, 6.88. As there are many named TCs not making landfall it is no surprise, that the range of all counterfactual histories is relatively big. While the best-case scenario only denotes 10 landfalls opposed to the observed 89, the worst-case scenario

assumes 187 landfalls for the same period. As such the influence of the counterfactual information leads to posterior beliefs suggesting much higher values for λ (Figure 4.3). Also, the posterior predictive distribution shows a large increase in the expected values for future landfall rates. The previously highest observed number of landfalls in a year is 10 landfalls. For the posterior predictive distribution $p(\tilde{y}|y_{obs})$ a value of 10 or more has a probability of 0.097. Based on the worst-case scenario, $p(\tilde{y}|y_{CFmax})$, the probability is 0.74 which is an increase of 64.44 %.

Also, for some individual countries and US states the differences between the worst-case scenario and the reference distribution are very big. For example, in Hawaii up to today never more than two landfalls in a year were reported and the historic posterior distribution shows a mean value of 0.71. This posterior value increases to 4.39 for the worst-case scenario. At the tail of the PP distribution the number of yearly landfalls at the 95th percentile increases from 2 to 8 TCs (Figure 4.6/Table A 3).

The data for US mainland needs to be viewed at with some caution. TCs usually move towards the west or north-west and away from the US mainland. This is the reason the TC risk is relatively low for continental US. However, at higher risk is Mexico, as EP TCs often form close to the Mexican coast. For Mexico the analysis of expected extreme landfall rates at the 95th percentile shows a considerable increase from 10 to 17 landfalls for the worst-case scenario compared to observed data. An active storm year with such high numbers of TC could thus lead to considerable losses and damages in Mexico.
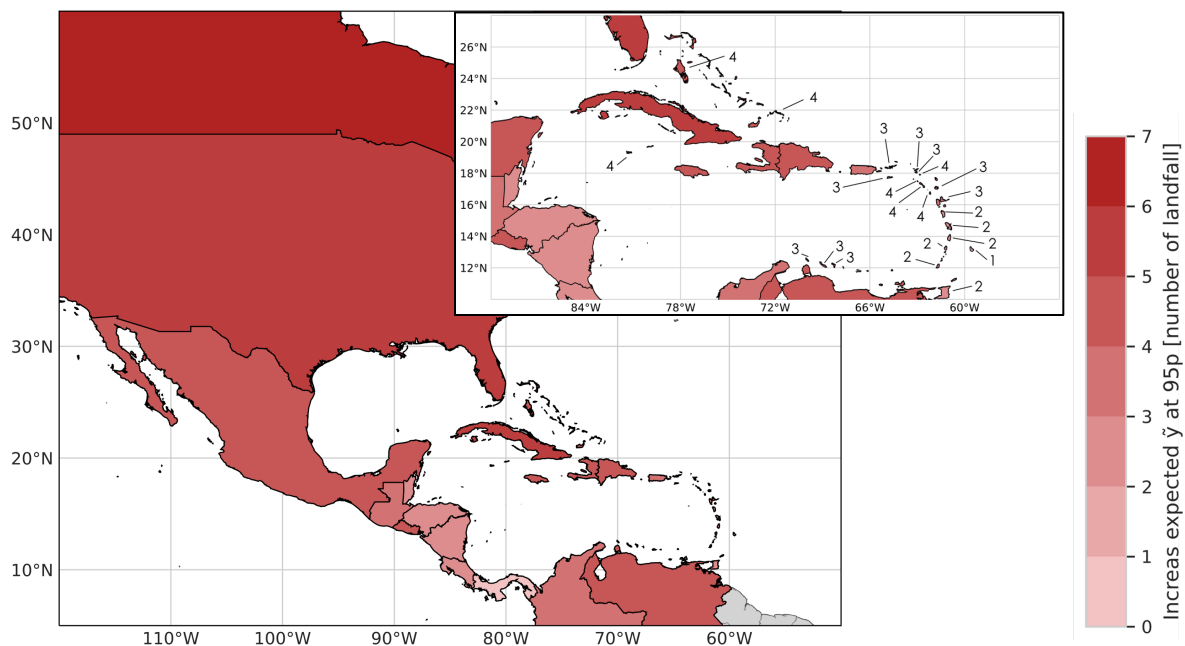

Figure 4.6: East Pacific - Difference in expected future landfall rates at 95th percentile of posterior predictive distribution, $E[\tilde{y}_{95p}|y_{CFmax}] - E[\tilde{y}_{95p}|y_{obs}]$. Relevant US states are shown individually while the rest of the US indicates the difference aggregated over the entire country.

West Pacific

The WP is the basin with the highest number of named TCs as well as the highest landfall rates. Most of all named events for the WP already make landfall (Figure 3.2, Table 3.1). In comparison, a counterfactual analysis of landfall rates does not reveal many worse-case

scenarios. The increase in belief for λ as well for future landfall rates is for the worst-case is the lowest across all basins in terms of percentages. The mean expected value for λ increases from 20.83 to 22.89 from the historic data to the belief based on the worst-case. Nevertheless, the probability of the previous maximum observed landfall rate of 28 or more to occur in the future increases by 7.2 % (Table 4.4).

Looking at the individual countries however, the picture changes immensely. Here the worst-case scenarios differ strongly form the BI based on observed TCs. For example, in Japan the expected mean landfall rate is almost twice as high based upon the worst-case counterfactual scenario compared to the posterior using observed landfall data. The posterior belief attributed to λ is 8.96 and 15.53 for $p(\lambda|y_{obs})$ and $p(\lambda|y_{CFmax})$ respectively (Table A 4). Also, the belief in future extreme landfall rates changes considerably based on downward counterfactuals. In Taiwan for example the PP distribution using observed TC tracks assumes 5 landfalls a year for a storm year at the 95th percentile. The worst-case scenario for Taiwan however would expect 15 landfalls within a year (Figure 4.7).
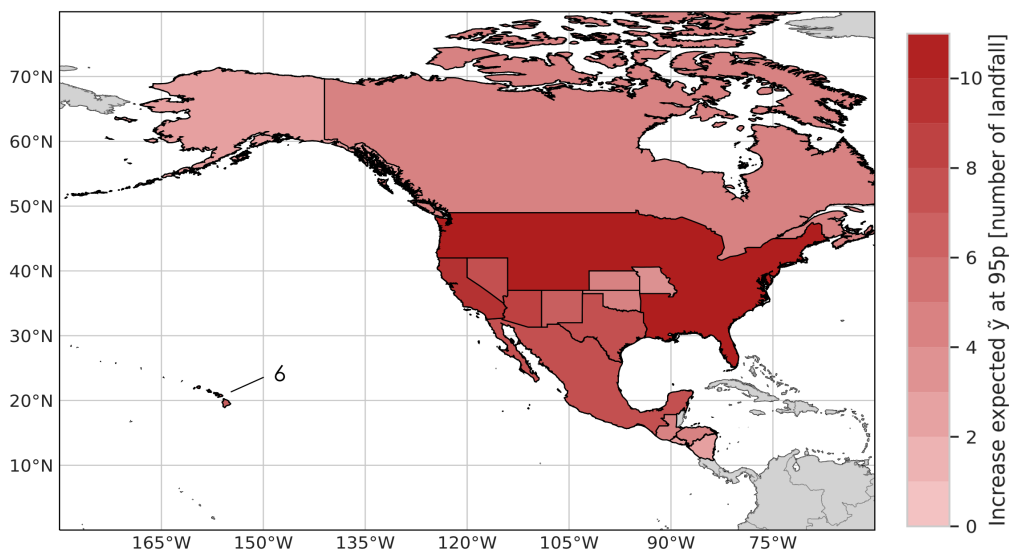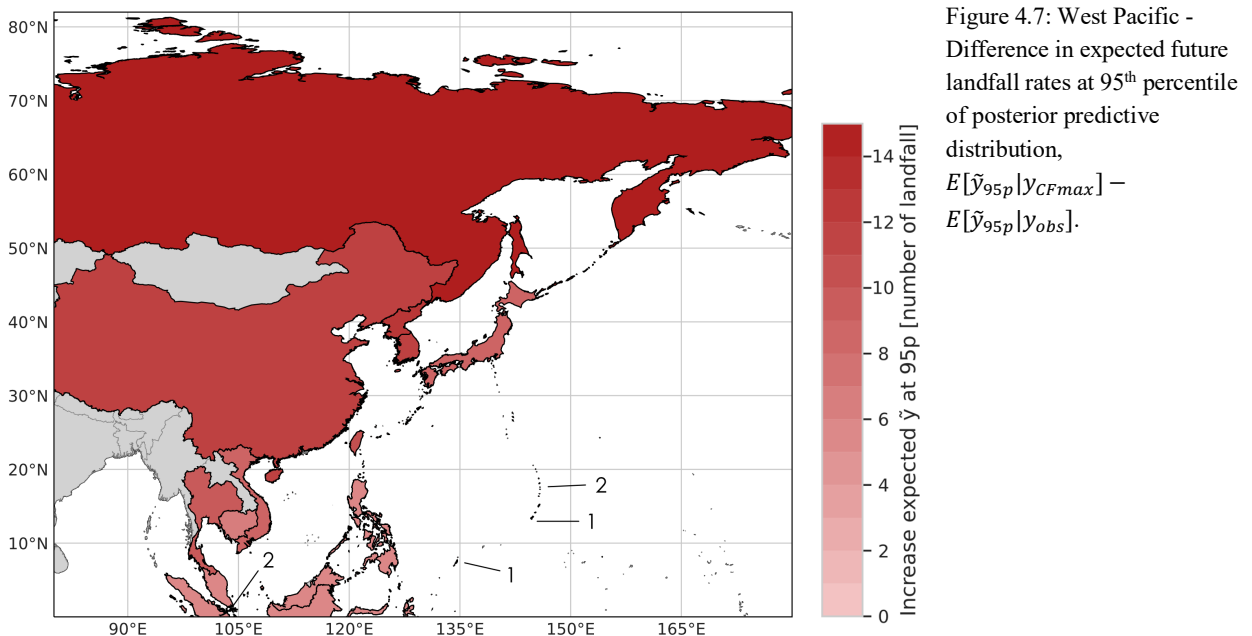

Figure 4.7: West Pacific - Difference in expected future landfall rates at 95th percentile of posterior predictive distribution, $E[\tilde{y}_{95p}|y_{CFmax}] - E[\tilde{y}_{95p}|y_{obs}]$.

North Indian Ocean

As in the WP basin in the NI most of the past TCs did make landfall. It makes therefore sense that the worst-case scenario with a sum of 60 landfalls does not by much exceed the observed 51 landfalls. The difference between the posterior distribution $p(\lambda|y_{CFmax})$ and $p(\lambda|y_{obs})$ is the smallest across all basins with $E[\lambda|y_{CFmax}] - E[\lambda|y_{obs}] = 0.56$. Also, the PP distribution show only little differences regarding expected future landfall rates (Figure 4.4). Up to 2007 the highest observed landfall rate is 10 ($ymax$). The probability of 10 of more landfalls in a year occurring in the future is very low for all observed and counterfactual scenarios. In the worst-case it only increases slightly. However, a landfall rate at or above the 95th percentile still increases from 8 to 9 landfalls within a year for the worst-case scenario.

In contrary to the results for the WP here the zoom into different countries does not reveal as many worse-case scenarios for the different countries either. Still for the different countries the worst-case scenarios do differ stronger from the posterior and poster predictive believes based

upon observed landfall rates. For all countries the belief in λ increases and the probability of observing the previous highest landfall rates in the future increases for some countries by up to 25 % (Table A 5). Future landfall rates at the 95th percentile for all countries increase by 1 or 2 landfalls a year when comparing the worst-case scenario to observations (Figure 4.8).

Figure 4.8: North Indian Ocean - Difference in expected future landfall rates at 95th percentile of posterior predictive distribution, $E[\tilde{y}_{95p}|y_{CFmax}] - E[\tilde{y}_{95p}|y_{obs}]$.

South Indian Ocean

The SI is more comparable to the EP as there are only few TCs making landfall (Figure 3.2). This again leaves room for a big spread of counterfactual scenarios. For the period 2008-2019 66 out of all 154 named TCs made landfall. By combining different forecast members however, the counterfactual scenarios show a range from 13 to 150 landfalls within the same period (Table 4.2). Percentage wise, the worst-case scenario in the SI shows the highest increase in the mean expected value for λ across all basins. The posterior belief based on observed TCs the mean landfall rate, $E[\lambda|y_{obs}]$, is at 5.79 landfalls a year. The worst-case scenario assumes almost twice as many with a mean value, $E[\lambda|y_{CFmax}]$, of 11.04. This is also expressed by a significant shift for the distribution of expected future LF rates (Figure 4.3). For example, the belief of observing the previously maximum observed landfall rate of 11 in the future increases by 40.6 % for the worst-case scenario. However, the PP distribution shows a much larger variance compared to other basins (Figure 4.4). This is due to the large difference of the counterfactual landfall rates to the prior information used for the Bayesian analysis.

Looking at individual countries the difference from the reference distribution to the counterfactual worst-case posterior and posterior predictive distribution are most distinct for small island states such as Mauritius or Réunion (Table A 6). This is especially true when looking at the tail end of the PP distribution. In Mauritius for example, a yearly landfall rate at the 95th percentile would be 2 landfalls based on observed values. In the worst-case scenario the PP distribution would assume 8 landfalls for the same percentile.
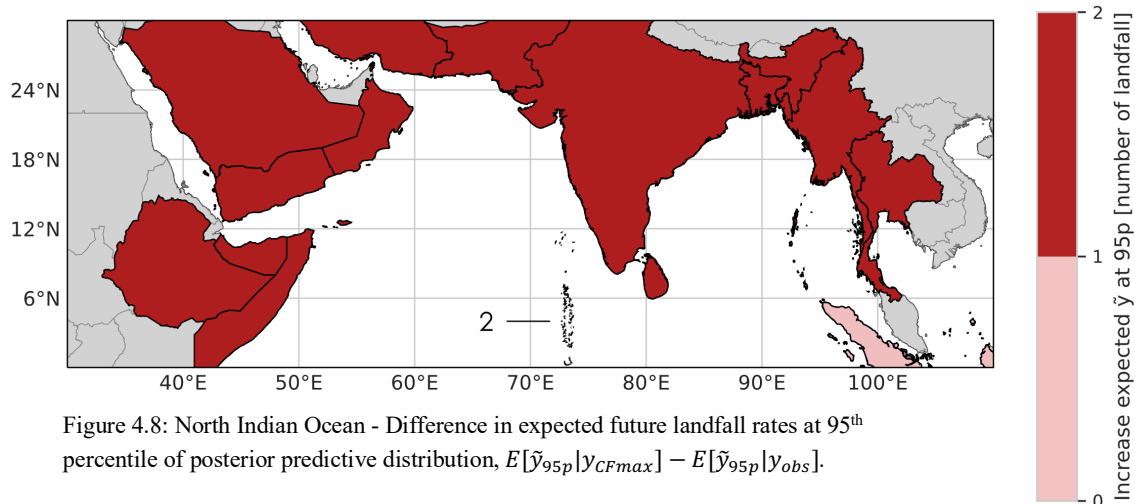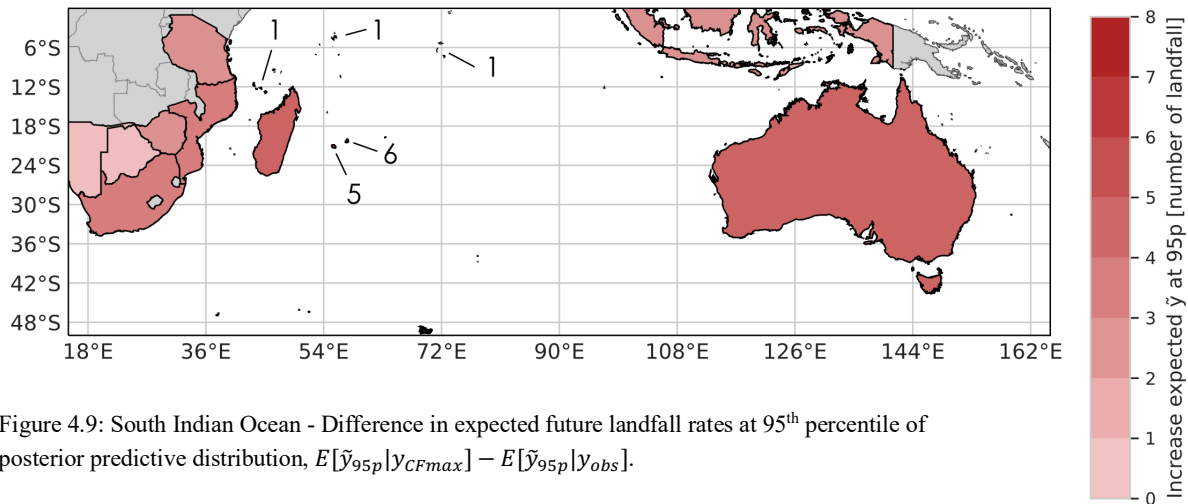
Figure 4.9: South Indian Ocean - Difference in expected future landfall rates at 95th percentile of posterior predictive distribution, $E[\tilde{y}_{95p}|y_{CFmax}] - E[\tilde{y}_{95p}|y_{obs}]$.

South Pacific

Within the SP basin 69 out of 101 occurring TCs between 2008 and 2019 made landfall. With around 68 % landfalls this is already a relatively high percentage. Still, the worst-case scenario shows a full 99 out of the 101 TCs making landfall (98 %). The resulting shift in posterior belief shows a difference of 1.88 landfalls a year form the posterior based on observed values to the worst-case scenario. The expected future values based upon the different posterior predictive distributions illustrate the same shift towards higher landfall rates. In the past the highest observed number of landfalls were 15 landfalls in 1998. Under the posterior predictive distribution based on observed data the probability of 15 or more landfalls is 0.09 %. The BI using the worst-case counterfactual scenario attributes the same values a probability of 0.96 % meaning an increase in probability by a factor of 10. However, as for the SI basin the posterior distribution, $p(\lambda|y_{obs})$, shows a shift towards lower landfall rates compared to the prior distribution. This should be considered when comparing the two distributions as the reference posterior might be undervaluing yearly landfall rates.

Looking at the individual countries the worst-case counterfactual scenarios differ much more from the observed values (Table A 7). For example, in Australia mean expected landfall rates increase by 3.44 landfalls per year $E[\lambda|y_{CFmax}] - E[\lambda|y_{obs}]$. Also, the probability of 5 landfalls within a year, the previous maximum observed value, or more increases by 47 %, almost doubling the belief in such extreme landfall rates. Similarly high shifts in probability towards higher landfall rates can be observed for most SP countries, including a few small island states. While the expected landfall rate at the 95th percentile increases by two for the entire basins, the difference for individual countries is up to 6 more landfalls in the worst-case scenario Figure 4.10. Very hard to interpret are islands where so far there were little to no landfalls were observed. For example, the first and only observed landfall in East Timor was by TC Gillian in 2014. However, the forecast data from 2008 to 2019 counts 11 more events with potential landfall and for the worst-case scenario even a maximum of 3 landfalls a year (Table A 7). Including those near-miss events we see a 10-fold increase in belief for mean yearly landfall rates in East Timor.
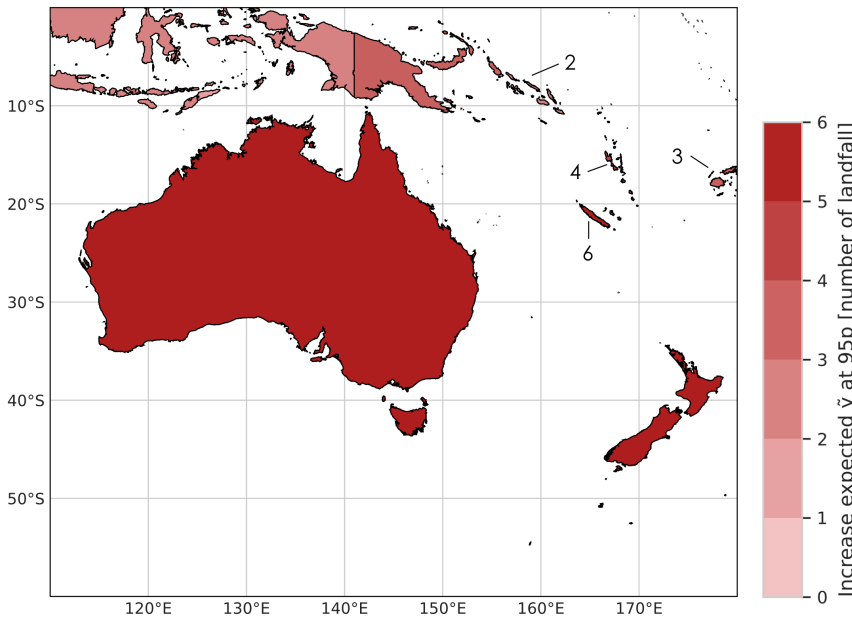
Figure 4.10: South Pacific - Difference in expected future landfall rates at 95th percentile of posterior predictive distribution, $E[\tilde{y}_{95p}|y_{CFmax}] - E[\tilde{y}_{95p}|y_{obs}]$.

4.3.2 Discussion across Basins

Focusing the attention on results based on observed landfall rates shows for all northern ocean basins an increase in belief for λ in the posterior distribution compared to the prior. For the two basins in the southern hemisphere however the trend is reversed. As discussed earlier an increase (decrease) in yearly TC events within the last two decades must be viewed with caution, considering changes in quality and availability of TC track data over the last few decades. Globally, positive trends in past landfall rates are mainly observed for intense TCs of category 3 or higher (IPCC, 2021). Statements concerning overall trend in frequencies can only be made with low confidence. Despite the increase(decrease) in landfall rates for the northern (southern) hemisphere no long-term trend should be assumed considering the short time span of the dataset. Especially with a negative trend in observed data a risk analysis based solely on past events might lead to an underestimation of future risk. Here a counterfactual risk analysis can be useful, as worse- and worst-case scenarios do reveal that higher landfall rates very well could have been observed given different circumstances.

Comparing the prior and posterior probability distribution of the observed landfall to the range of all counterfactual scenarios puts both within inside the range from best- to worst-case scenarios for all basins. This is important to note, as it means that the different combinations of forecast tracks do indeed allow for the creation of upward and downward counterfactual scenarios for yearly landfall rates. The differences in range across basins when comparing counterfactual scenarios to observed landfall rates can have several reasons. For one, not all basins have the same number of TCs during a year or season (Table 3.1). The period 2008 to 2019 counted 288 tropical storms and TCs in the WP basin, while the NI basin only counted 65. Another explanation can be due to the geographic situation. The NA, WP and NI basins showing lower differences from the observed landfall rates to the worst-case do have larger bodies of land and thus more coastlines. Especially the NI ocean is relatively small compared to other basins and pretty much encircled by land. Thus, it is of no surprise that most occurring TCs do make landfall and that there is not a much higher potential in counterfactual analysis.

The other extreme is the EP basin, where most TCs never make landfall and the counties most at risk are relatively small island states. Finally, differences could also be due to different data availability of forecasts. Especially in the NI basin where also non-named TCs are included, there are larger gaps in the forecast dataset (Table 3.2).

In summary, the range of the counterfactual scenarios does depend mostly on the geographic conditions and the percentage of named TCs making landfall. Especially for the NI and WP, where most observed TCs make landfall there are not as many near-miss events that could contribute to a downward counterfactual scenario. The same holds true to some degree for the WP, NA, and SP basin. On the other hand, for the WP and SI basin there is a large number of near-miss events, leading to a much larger range and more downward counterfactual scenarios. The worst-case scenarios show much larger deviations in posterior and posterior predictive distributions from the reference distribution based on observation data. Still, for all basins there are counterfactual scenarios indicating potentially higher landfall rates than what previously was observed. Based on the worst-case scenario mean landfall rates and future landfall rates would be expected to be higher and extreme landfall rates more extreme in all case studies. Such high landfall rates at the tail of the distribution might not be included in the risk horizon when assessing risk solely on observational data. For the basins where there are not many downward counterfactual scenarios on simple landfall counts it might be more interesting to look further into intensity variable of past TC, such as central pressure or wind speed. For those the incorporation of forecast data in counterfactual risk analysis might change the posterior belief and thus a future risk assessment considerably.

Also, focusing the analysis of landfall rates on a smaller scale can reveal more potential for (downward) counterfactual scenarios. On this scale the different paths a TC could have taken strongly influences the potential impacts of the surrounding countries of a basin. Looking at the results for the individual countries reveals a more detailed picture of the downward counterfactual potential. In all basins the worst-case scenario does lead to higher deviations from observed landfalls than when aggregating over the entire basin. Especially for the WP where the analysis across the basin does not lead to a higher attribution of extreme landfall rates, some countries do reveal large potential for higher landfall rates. This is impressively demonstrated for example in Japan or Taiwan but also other states of the region. On the other hand, in the NI basin also the zoom into the different countries does not reveal such a strong potential for downward counterfactuals. Even more than the WP the NI ocean is surrounded by land and TCs missing land are rather rare. Compared to the WP the countries and thus the size of the coastline of the individual countries is bigger. Additionally, by far the biggest part of the coastline is in India. So maybe in this case it might be interesting to zoom in even more and look for individual cities situated at the coastline. Then different paths an event could have taken can be better distinguished while now they just might be aggregated in Indian landfalls. However, another reason for the small differences could also be, that for many events in period 2008-2019 there are no forecast data available, as not named TCs had to be included.

Across all basins very interesting are the results for small islands states, where past TC landfalls are very rare and there is little data on which the risk due to TCs can be computed. Incorporating forecast tracks in the analysis shows that for all countries there are some near-miss events. Incorporating those near-miss events can significantly improve the data availability when

assessing TC risk and increase the assessment of potential landfall rates. As an example, the risk assessment for Mauritius could be very different when an extreme year concerning landfalls in Mauritius is thought of as 2 (observation data) or 8 landfalls (worst-case scenario). This might also influence decisions when considering mitigation measures. Another example is the EP basin which in general is not known for high impacts due to TCs. This is also illustrated by the fact that only few of the named TCs actually make landfall. However, there are some small islands at risk of TC hazards with potentially large losses and damages. This analysis shows very impressively how devastating a worse or even worst-case scenario could be for example for Hawaii. On the other hand, for a big country such as the United States or Australia the simple number of landfalls might not be so relevant as the exact place the TC hits the coast. As the losses and damages depend strongly on whether the TC hits a big city or makes landfall in an uninhabited part of the country. For this reason, it might be interesting to combine an analysis of counterfactual track data with data on population density and exposure in future projects.

5   Conclusions and Outlook

Within this thesis Bayesian inference is used to incorporate counterfactual information based on forecast tracks of past TC events, in a risk analysis of global TC landfall rates. To assess the sensitivity of the analysis different periods to build the prior belief and datasets from different providers of forecast tracks are compared.

By combining different forecast tracks of past TC events any number of counterfactual scenarios for TC landfall rates can be built. The manually created scenarios are based on the probability of a TC making landfall considering its forecast range. This allows for the creation of upward and downward counterfactual scenarios that envelope the observed landfall rates. To improve risk analysis of extreme hazard events, forecast data also allows for the specific search for near-miss events and downward counterfactual scenarios of past landfall rates. Limitations may arise due to the quality and resolution of the forecast data that differs in time, region and across providers. Since the presented analysis only includes named events, the results are somewhat biased towards these events. The density of available TC data for risk analysis might even be further expanded when also considering forecast data for non-named or invest events. Even so, the analysis shows that forecast data can be a powerful source for the creation of counterfactual scenarios of past TC events. For all basins and countries, the analysis of downward counterfactual scenarios leads to a higher estimation of expected mean yearly landfall rates, probability of extreme events and outliers as well as future extreme yearly landfall rates.

The findings across case studies including different observation periods for the prior distribution does not show large differences for most basins. Also, analysis based on different forecast providers in the creation of counterfactual scenarios might not show strong sensitivities for the BI based risk analysis. However, they do highlight the importance of dense TC track dataset of good quality to achieve a risk analysis accounting for extreme hazard scenarios. By only including observation data from the modern satellite era starting 1980 the number of missed storms is reduced and the dataset is comparable across all basins. Data quality is also a concern when including forecast data in a counterfactual risk analysis. There are differences in

data resolution, model parametrizations and data availability across providers, period, and regions. But while for observation data it is important to have a comparable and homogeneous dataset in counterfactual risk analysis the focus is rather on exploring the realm of possible scenarios. So here it makes sense to combine forecast data from different providers. By doing so the range of the counterfactual scenarios can be augmented and the amount of data gaps is reduced.

Findings across basins and countries show large differences in the range of counterfactual scenarios and the potential for downward counterfactuals based on near-miss events. For basins with lower percentage of past events making landfall the range of counterfactual landfall rates is larger, and the possible worst-case scenario differs stronger from the observed landfall rates. The results aggregated over the entire basin however does not necessarily reflect the risk of individual countries. Looking at individual countries, counterfactual risk analysis of landfall rates reveals a larger capacity for downward counterfactual scenarios. Especially for small island states including such near-miss events in risk analysis can have a large effect in the assessment of potential losses and damages. For small islands usually the number of observed landfalls is very small. Such sparse data availability makes a correct assessment of extreme hazard scenarios difficult. Thus, counterfactual information is able to increase the data availability and improve the detection of potential extremes. Also, for bigger countries counterfactual data can have a large impact on the risk assessment. However, for future works it would be interesting to look at TC landfalls at a higher spatial resolution on a regional or local scale.

The presented analysis shows how extreme numbers of yearly landfall rates might look like if incorporating downward counterfactual data in risk analysis. However, no concluding statements can be made regarding potential losses of damages. Further counterfactual risk analysis of TC is therefore necessary where TC tracks are combined with information on exposure and vulnerability. Also, of high relevance for when assessing hazard risk is the intensity of a storm. Future counterfactual risk analysis might therefore include more TC variables such as sustained wind speed, central pressure, or translation speed. However, for a counterfactual analysis of pressure or wind data from forecast data, a more thorough quality control and correction of biases based on the resolution of forecast data would be necessary. As a first exploration of counterfactual risk analysis of tropical cyclones based on forecast data the presented results are promising. Forecast data allows for the detection of near-miss events and the creation of downward counterfactual scenarios. These can be useful in improving risk assessment of tropical cyclones at the tail of the probability distribution.

References

Aspinall, W. and Woo, G. (2019) 'Counterfactual Analysis of Runaway Volcanic Explosions', *Frontiers in Earth Science*, 7(August), pp. 1–15.

Bassill, N. P. (2014) 'Accuracy of early GFS and ECMWF Sandy (2012) track forecasts: Evidence for a dependence on cumulus parameterization', *Geophysical Research Letters*, 41(9), pp. 3274–3281.

Bertogg, M. (2021) *Why hurricane risk modelling has to change, Swiss Re*. Available at:

https://www.swissre.com/risk-knowledge/mitigating-climate-risk/why-hurricane-risk-modelling-has-to-change.html (Accessed: 7 January 2022).

Bloemendaal, N. *et al.* (2020) 'Generation of a global synthetic tropical cyclone hazard dataset using STORM', *Scientific Data*. Springer US, 7(1), pp. 1–12.

Bougeault, P. *et al.* (2010) 'The thorpex interactive grand global ensemble', *Bulletin of the American Meteorological Society*, 91(8), pp. 1059–1072.

Britannica, The Editors of Encyclopaedia (2021) *Hurricane Katrina*, *Encyclopedia Britannica*. Available at: https://www.britannica.com/event/Hurricane-Katrina (Accessed: 16 January 2022).

Donovan, T. M. and Mickey, R. M. (2019) *Bayesian Statistics for Beginners: a step-by-step approach*. 1st edn. Oxford: Oxford University Press, USA. Available at: https://oxford.universitypressscholarship.com/view/10.1093/oso/9780198841296.001.0001/oso-9780198841296-miscMatter-1.

Elsner, J. B. and Bossak, B. H. (2001) 'Bayesian analysis of U.S. hurricane climate', *Journal of Climate*, 14(23), pp. 4341–4350.

Gelman, A. *et al.* (2020) *Bayesian Data Analysis*. Third edition (with errors fixed as of 13 February 2020): electronic edition.

Gill, J. (2015) *Bayesian Methods. A Social and Behavioral Sciences Approach*. 3rd Editio. New York: Chapman and Hall/CRC.

IPCC (2014) 'Summary for policy makers', in Field, C. B. et al. (eds) *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, pp. 1–32.

IPCC (2021) 'Summary for Policymakers', in Masson- Delmotte, V. et al. (eds) *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. In Press.

Knapp, K. R. *et al.* (2010) 'The international best track archive for climate stewardship (IBTrACS)', *Bulletin of the American Meteorological Society*, 91(3), pp. 363–376.

Knapp, K. R. *et al.* (2018) *International Best Track Archive for Climate Stewardship (IBTrACS) Project, Version 4.*, *NOAA National Centers for Environmental Information*. Available at: https://doi.org/10.25921/82ty-9e16 (Accessed: 29 September 2021).

Knapp, K. R. (2019) 'International Best Track Archive for Climate Stewardship (IBTrACS). Technical documentation', *National Oceanic and Atmospheric Administration, National Climatic Data Center*, pp. 1–24. Available at: https://www.ncdc.noaa.gov/ibtracs/pdf/IBTrACS_version4_Technical_Details.pdf%0Ahttps://www.ncdc.noaa.gov/ibtracs/index.php.

Lackmann, G. M. (2015) 'Hurricane Sandy before 1900 and after 2100', *Bulletin of the American Meteorological Society*, 96(4), pp. 547–560.

Lohmann, U., Lüönd, F. and Mahrt, F. (2016) *An introduction to clouds: From the microscale*

to climate. Cambridge University Press.

Munich RE (2017) 'A stormy year TOPICS Geo Natural catastrophes 2017', *TOPICS geo*.

National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce *et al.* (2008) *THORPEX Interactive Grand Global Ensemble (TIGGE) Model Tropical Cyclone Track Data*. Boulder, CO: Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. Available at: https://doi.org/10.5065/D6GH9GSZ (Accessed: 2 April 2020).

NOAA, National Hurricane Center and Central Pacific Hurricane Center (2021) *Saffir-Simpson Hurricane Wind Scale*. Available at: https://www.nhc.noaa.gov/aboutsshws.php (Accessed: 6 January 2022).

Oughton, E. J. *et al.* (2019) 'Stochastic Counterfactual Risk Analysis for the Vulnerability Assessment of Cyber-Physical Attacks on Electricity Distribution Infrastructure Networks', *Risk Analysis*, 39(9), pp. 2012–2031.

Philip, T. *et al.* (2019) 'Issues of Importance to the (Re)insurance Industry: A Timescale Perspective', in Collins, J. M. and Walsh, K. (eds) *Hurricane Risk*. 1st edn. Cham, Switzerland: Springer, pp. 1–22.

Rappaport, E. N. *et al.* (2009) 'Advances and challenges at the national hurricane center', *Weather and Forecasting*, 24(2), pp. 395–419.

Schreck, C. J., Knapp, K. R. and Kossin, J. P. (2014) 'The impact of best track discrepancies on global tropical cyclone climatologies using IBTrACS', *Monthly Weather Review*, 142(10), pp. 3881–3899.

The CTHAEH (2016) *Frequentist and Bayesian Approaches in Statistic*, *Probabilistic World*. Available at: https://www.probabilisticworld.com/frequentist-bayesian-approaches-inferential-statistics/ (Accessed: 25 May 2020).

Wikipedia (2021) *Saffir-Simpson scale*. Available at: https://en.wikipedia.org/wiki/Saffir–Simpson_scale (Accessed: 6 January 2022).

Woo, G. (2016) 'Counterfactual Disaster Risk Analysis', *Variance*, 10(2), pp. 1–30.

Woo, G. (2019) 'Downward Counterfactual Search for Extreme Events', *Frontiers in Earth Science*, 7(December).

Woo, G., Maynard, T. and Seria, J. (2017) 'Reimagining history: Counterfactual risk analysis', p. 49.

Yamaguchi, M., Nakazawa, T. and Hoshino, S. (2012) 'On the relative benefits of a multi-centre grand ensemble for tropical cyclone track prediction in the western North Pacific', *Quarterly Journal of the Royal Meteorological Society*, 138(669), pp. 2019–2029.

Zehnder, J. A. (2021) *tropical cyclone*, *Encyclopedia Britannica*. Available at: https://www.swissre.com/risk-knowledge/mitigating-climate-risk/why-hurricane-risk-modelling-has-to-change.html (Accessed: 28 January 2022).

Appendix

Appendix A   Basins – Extended Results

Table A 1: Prior distribution, posterior distribution, and posterior predictive distribution from observed data and counterfactual data. Comparison across different prior periods and forecast data providers.

Basin	Period prior	Provider	Mean prior μ_{hist}	Mean posterior observed $E[\lambda\|y_{obs}]$	Mean posterior worst-case CF $E[\lambda\|y_{CFmax}]$	Difference posterior $E[\lambda\|y_{CFmax}] - E[\lambda\|y_{obs}]$	Max. observed y 80-07 $(ymax)$	$p(\tilde{y} \geq ymax\|y_{obs})$	$p(\tilde{y} \geq ymax\|y_{CFmax})$	$\Delta p(\tilde{y} \geq ymax)$ [%]	$E[\tilde{y}_{95p}\|y_{obs}]$	$E[\tilde{y}_{95p}\|y_{CF_{maxLF}}]$
EP	1950 - 2007	ECMWF	4.824	6.768	12.143	5.375	10	0.0892	0.6610	57.18%	11	18
		KWBC			12.268	5.500			0.6734	58.43%		18
		KWBC+ECMWF			12.893	6.125			0.7313	64.22%		19
	1880 - 2007	ECMWF	5.286	6.884	12.259	5.375	10	0.0968	0.6725	57.56%	12	18
		KWBC			12.384	5.500			0.6847	58.78%		19
		KWBC+ECMWF			13.009	6.125			0.7413	64.44%		19
NA	1950 - 2007	ECMWF	6.207	7.927	9.677	1.750	17	0.0021	0.0130	1.09%	13	15
		KWBC			9.802	1.875			0.0145	1.24%		15
		KWBC+ECMWF			10.302	2.375			0.0219	1.98%		16
	1880 - 2007	ECMWF	6.714	8.054	9.804	1.750	17	0.0024	0.0145	1.21%	13	15
		KWBC			9.929	1.875			0.0161	1.37%		16
		KWBC+ECMWF			10.429	2.375			0.0241	2.17%		16
NI	1950 - 2007	ECMWF	7.379	5.032	5.470	0.438	17	0.0000	0.0000	0.00%	9	10
		KWBC			5.282	0.250			0.0000	0.00%		9
		KWBC+ECMWF			5.595	0.563			0.0000	0.00%		10
	1880 - 2007	ECMWF	4.071	4.205	4.643	0.438	10	0.0055	0.0104	0.50%	8	9
		KWBC			4.455	0.250			0.0080	0.25%		8
		KWBC+ECMWF			4.768	0.563			0.0123	0.69%		9

	Period	Model									
SI	1951 - 2008	ECMWF			9.652	3.813			0.0725	7.10%		15
		KWBC	6.857	5.839	10.589	4.750	14	0.0016	0.1244	12.28%	10	16
		KWBC+ECMWF			11.089	5.250			0.1590	15.75%		17
	1881 - 2008	ECMWF			9.598	3.813			0.2631	24.43%		15
		KWBC	6.643	5.786	10.536	4.750	11	0.0188	0.3671	34.83%	10	16
		KWBC+ECMWF			11.036	5.250			0.4248	40.61%		17
SP	1951 - 2008	ECMWF			7.337	1.438			0.0050	0.43%		12
		KWBC	6.347	5.899	7.587	1.688	15	0.0007	0.0066	0.59%	10	13
		KWBC+ECMWF			7.774	1.875			0.0081	0.74%		13
	1881 - 2008	ECMWF			7.491	1.438			0.0060	0.51%		12
		KWBC	6.964	6.054	7.741	1.688	15	0.0009	0.0078	0.70%	11	13
		KWBC+ECMWF			7.929	1.875			0.0096	0.87%		13
WP	1950 - 2007	ECMWF			22.610	1.938			0.0595	3.60%		31
		KWBC	20.190	20.672	22.610	1.938	30	0.0235	0.0595	3.60%	29	31
		KWBC+ECMWF			22.735	2.063			0.0627	3.92%		31
	1880 - 2007	ECMWF			22.768	1.938			0.1240	6.64%		31
		KWBC	20.821	20.830	22.768	1.938	28	0.0575	0.1240	6.64%	29	31
		KWBC+ECMWF			22.893	2.063			0.1294	7.19%		31

Appendix B   Countries – Results Prior, Posterior and Posterior Predictive Distribution

Table A 2: North Atlantic - Prior distribution, posterior distribution, and posterior predictive distribution from observed data and counterfactual data.

Country	Mean prior μ_{hist}	Mean posterior observed $E[\lambda|y_{obs}]$	Mean posterior worst-case CF $E[\lambda|y_{CFmax}]$	Difference posterior $E[\lambda|y_{CFmax}] - E[\lambda|y_{obs}]$	Max. observed y 80-07 ($ymax$)	$p(\tilde{y} \geq ymax| y_{obs})$	$p(\tilde{y} \geq ymax| y_{CFmax})$	$\Delta p(\tilde{y} \geq ymax)$ [%]	$E[\tilde{y}_{95p}|y_{obs}]$	$E[\tilde{y}_{95p}| y_{CFmax}]$
Anguilla	0.321	0.455	2.330	1.875	2	0.014	0.409	39.50%	2	5
Antigua and Barbuda	0.357	0.464	2.214	1.750	2	0.015	0.379	36.43%	2	5
Aruba	0.071	0.018	0.955	0.938	1	0.001	0.248	24.75%	0	3
Bahamas	1.179	1.545	4.295	2.750	5	0.007	0.266	25.90%	4	8
Barbados	0.250	0.500	1.063	0.563	3	0.003	0.027	2.43%	2	3
Belize	0.286	0.821	2.009	1.188	2	0.056	0.326	26.98%	3	5
Bonaire	0.107	0.027	0.839	0.813	1	0.001	0.207	20.59%	0	3
British Virgin Islands	0.357	0.402	2.464	2.063	2	0.011	0.443	43.26%	2	5
Canada	1.429	1.545	6.357	4.813	4	0.025	0.751	72.68%	4	11
Cayman Islands	0.429	0.232	2.295	2.063	3	0.000	0.204	20.35%	1	5
Colombia	0.250	0.250	1.875	1.625	3	0.000	0.127	12.63%	1	4
Costa Rica	0.036	0.071	1.134	1.063	1	0.004	0.312	30.76%	1	3
Cuba	0.964	1.179	4.179	3.000	3	0.037	0.594	55.74%	3	8
Curaçao	0.107	0.027	1.089	1.063	1	0.001	0.296	29.49%	0	3
Dominica	0.179	0.482	1.670	1.188	2	0.016	0.237	22.09%	2	4
Dominican Republic	0.536	0.509	3.009	2.500	2	0.019	0.572	55.34%	2	6
El Salvador	0.107	0.027	1.339	1.313	1	0.001	0.383	38.23%	0	4
Grenada	0.464	0.179	1.054	0.875	3	0.000	0.027	2.64%	1	3
Guadeloupe	0.393	0.598	1.973	1.375	2	0.027	0.316	28.91%	2	5
Guatemala	0.179	0.607	2.170	1.563	1	0.128	0.629	50.12%	2	5

Haiti	0.429	0.607	2.857	2.250	2	0.028	0.538	50.99%	2	6
Honduras	0.714	1.116	2.366	1.250	3	0.031	0.218	18.67%	3	5
Jamaica	0.429	0.232	2.232	2.000	2	0.003	0.384	38.10%	1	5
Martinique	0.250	0.375	1.625	1.250	2	0.009	0.226	21.68%	2	4
Mexico	1.500	2.063	4.625	2.563	7	0.002	0.104	10.15%	5	9
Montserrat	0.250	0.313	2.000	1.688	1	0.044	0.586	54.15%	1	5
Nicaragua	0.500	0.625	1.875	1.250	2	0.030	0.290	26.05%	2	4
Panama	0.000	0.000	1.063	1.063	0	0.000	0.643	64.32%	0	3
Puerto Rico	0.500	0.563	2.438	1.875	2	0.024	0.437	41.31%	2	5
Saba	0.179	0.170	1.920	1.750	2	0.001	0.302	30.07%	1	5
Saint Kitts and Nevis	0.179	0.232	1.982	1.750	2	0.003	0.319	31.55%	1	5
Saint Lucia	0.250	0.500	1.688	1.188	1	0.095	0.496	40.18%	2	4
Saint Vincent and the Grenadines	0.357	0.527	1.402	0.875	3	0.003	0.059	5.60%	2	4
Saint-Barthélemy	0.286	0.259	2.259	2.000	2	0.004	0.391	38.71%	1	5
Saint-Martin	0.286	0.384	2.259	1.875	2	0.010	0.391	38.13%	2	5
Sint Eustatius	0.179	0.170	1.920	1.750	2	0.001	0.302	30.07%	1	5
Trinidad and Tobago	0.321	0.143	0.893	0.750	3	0.000	0.016	1.62%	1	3
Turks and Caicos Islands	0.321	0.705	2.893	2.188	2	0.040	0.546	50.66%	2	6
United States	3.857	3.527	7.527	4.000	9	0.005	0.232	22.70%	7	12
United States Virgin Islands	0.500	0.500	2.375	1.875	2	0.018	0.421	40.29%	2	5
Venezuela	0.250	0.313	2.250	1.938	2	0.006	0.389	38.26%	1	5

Table A 3: East Pacific - Prior distribution, posterior distribution, and posterior predictive distribution from observed data and counterfactual data.

Country	Mean prior μ_{hist}	Mean posterior observed $E[\lambda|y_{obs}]$	Mean posterior worst-case CF $E[\lambda|y_{CFmax}]$	Difference posterior $E[\lambda|y_{CFmax}] - E[\lambda|y_{obs}]$	Max. observed y 80-07 (ymax)	$p(\widetilde{y} \geq ymax| y_{obs})$	$p(\widetilde{y} \geq ymax| y_{CFmax})$	$\Delta p(\widetilde{y} \geq ymax)$ [%]	$E[\widetilde{y}_{95p}|y_{obs}]$	$E[\widetilde{y}_{95p}| y_{CFmax}]$
Alaska	0.000	0.063	1.313	1.250	0	0.059	0.720	66.12%	1	3
Arizona	0.107	0.089	5.152	5.063	1	0.006	0.960	95.46%	1	9
California	0.036	0.009	4.821	4.813	1	0.000	0.948	94.82%	0	9
Canada	0.000	0.063	2.188	2.125	0	0.059	0.880	82.14%	1	5
El Salvador	0.107	0.152	1.089	0.938	1	0.014	0.296	28.24%	1	3
Guatemala	0.071	0.080	2.268	2.188	1	0.005	0.653	64.76%	1	5
Hawaii	0.321	0.705	4.393	3.688	2	0.040	0.805	76.57%	2	8
Honduras	0.071	0.080	1.955	1.875	1	0.005	0.574	56.87%	1	5
Kansas	0.000	0.063	2.063	2.000	0	0.059	0.865	80.59%	1	5
Mexico	4.893	5.973	11.223	5.250	9	0.088	0.676	58.79%	10	17
Missouri	0.000	0.063	1.875	1.813	0	0.059	0.838	77.89%	1	4
Nevada	0.036	0.009	3.446	3.438	1	0.000	0.850	84.98%	0	7
New Mexico	0.071	0.018	3.018	3.000	1	0.001	0.794	79.37%	0	6
Nicaragua	0.036	0.071	1.009	0.938	1	0.004	0.267	26.31%	1	3
Oklahoma	0.000	0.063	2.500	2.438	0	0.059	0.912	85.27%	1	5
Texas	0.071	0.080	3.955	3.875	1	0.005	0.898	89.31%	1	8
United States	0.500	0.938	9.000	8.063	3	0.019	0.976	95.71%	3	14

Table A 4: West Pacific - Prior distribution, posterior distribution, and posterior predictive distribution from observed data and counterfactual data.

Country	Mean prior μ_{hist}	Mean posterior observed $E[\lambda|y_{obs}]$	Mean posterior worst-case CF $E[\lambda|y_{CFmax}]$	Difference posterior $E[\lambda|y_{CFmax}] - E[\lambda|y_{obs}]$	Max. observed y 80-07 (ymax)	$p(\tilde{y} \geq ymax| y_{obs})$	$p(\tilde{y} \geq ymax| y_{CFmax})$	$\Delta p(\tilde{y} \geq ymax)$ [%]	$E[\tilde{y}_{95p}|y_{obs}]$	$E[\tilde{y}_{95p}| y_{CFmax}]$
Cambodia	0.786	0.821	4.759	3.938	4	0.002	0.512	50.95%	3	9
China	8.179	8.607	17.295	8.688	13	0.061	0.810	74.85%	14	25
Dem. Rep. Korea	0.250	0.625	8.563	7.938	1	0.134	0.998	86.40%	2	14
Guam	0.679	0.920	1.607	0.688	3	0.018	0.086	6.78%	3	4
Indonesia	0.107	0.027	2.527	2.500	1	0.001	0.709	70.77%	0	5
Japan	8.857	8.964	15.527	6.562	16	0.013	0.389	37.54%	14	22
Malaysia	0.071	0.268	2.705	2.438	1	0.034	0.743	70.85%	1	6
Northern Mariana Islands	3.679	3.107	4.420	1.313	10	0.001	0.008	0.69%	6	8
Palau	0.893	0.786	1.348	0.563	4	0.002	0.015	1.32%	3	4
Philippines	7.929	7.732	11.545	3.813	13	0.031	0.276	24.48%	13	18
Republic of Korea	1.500	1.625	9.250	7.625	4	0.029	0.948	91.88%	4	15
Russian Federation	0.536	0.259	10.384	10.125	2	0.004	0.998	99.37%	1	16
Singapore	0.036	0.009	0.759	0.750	1	0.000	0.179	17.86%	0	2
Taiwan	2.786	2.446	9.571	7.125	5	0.043	0.908	86.48%	5	15
Thailand	1.143	1.286	7.098	5.813	4	0.013	0.827	81.46%	3	12
Vietnam	4.071	4.518	10.643	6.125	8	0.046	0.727	68.08%	8	16

Table A 5: North Indian Ocean - Prior distribution, posterior distribution, and posterior predictive distribution from observed data and counterfactual data.

| Country	Mean prior μ_{hist}	Mean posterior observed $E[\lambda|y_{obs}]$	Mean posterior worst-case CF $E[\lambda|y_{CFmax}]$	Difference posterior $E[\lambda|y_{CFmax}] - E[\lambda|y_{obs}]$	Max. observed y 80-07 ($ymax$)	$p(\tilde{y} \geq ymax | y_{obs})$	$p(\tilde{y} \geq ymax | y_{CFmax})$	$\Delta p(\tilde{y} \geq ymax)$ [%]	$E[\tilde{y}_{95p}|y_{obs}]$	$E[\tilde{y}_{95p}|y_{CFmax}]$
Bangladesh	0.893	0.786	1.911	1.125	2	0.050	0.300	24.93%	3	5
Ethiopia	0.036	0.071	0.821	0.750	1	0.004	0.201	19.65%	1	3
India	3.179	2.920	3.920	1.000	7	0.013	0.052	3.90%	6	8
Indonesia	0.071	0.018	0.018	0.000	1	0.001	0.001	0.00%	0	0
Iran	0.071	0.018	0.205	0.188	2	0.000	0.002	0.22%	0	1
Maldives	0.179	0.045	0.670	0.625	2	0.000	0.035	3.50%	0	2
Myanmar	0.929	0.732	1.295	0.563	3	0.009	0.048	3.87%	2	3
Oman	0.250	0.313	1.000	0.688	1	0.044	0.264	21.98%	1	3
Pakistan	0.179	0.107	0.857	0.750	1	0.008	0.213	20.55%	1	3
Saudi Arabia	0.107	0.089	0.902	0.813	1	0.006	0.229	22.32%	1	3
Somalia	0.179	0.357	1.107	0.750	1	0.055	0.302	24.72%	2	3
Somaliland	0.107	0.214	0.964	0.750	1	0.024	0.251	22.75%	1	3
Sri Lanka	0.321	0.455	1.080	0.625	2	0.014	0.101	8.67%	2	3
Thailand	0.143	0.098	0.411	0.313	1	0.007	0.069	6.22%	1	2
Yemen	0.143	0.473	1.161	0.688	1	0.087	0.321	23.46%	2	3

Table A 6: South Indian Ocean - Prior distribution, posterior distribution, and posterior predictive distribution from observed data and counterfactual data.

Country	Mean prior μ_{hist}	Mean posterior observed $E[\lambda|y_{obs}]$	Mean posterior worst-case CF $E[\lambda|y_{CFmax}]$	Difference posterior $E[\lambda|y_{CFmax}] - E[\lambda|y_{obs}]$	Max. observed y 80-07 (ymax)	$p(\tilde{y} \geq ymax| y_{obs})$	$p(\tilde{y} \geq ymax| y_{CFmax})$	$\Delta p(\tilde{y} \geq ymax)$ [%]	$E[\tilde{y}_{95p}|y_{obs}]$	$E[\tilde{y}_{95p}| y_{CFmax}]$
Australia	2.429	2.107	4.857	2.750	6	0.008	0.222	21.42%	5	9
Botswana	0.000	0.000	0.563	0.563	0	0.000	0.421	42.05%	0	2
British Indian Ocean Territory	0.071	0.080	0.393	0.313	1	0.005	0.064	5.93%	1	2
Comoros	0.143	0.161	0.723	0.563	1	0.015	0.167	15.17%	1	2
French Southern and Antarctic Lands	1.036	0.884	6.196	5.313	4	0.003	0.732	72.88%	3	11
Indian Ocean Territories	0.357	0.214	1.027	0.813	2	0.003	0.091	8.82%	1	3
Indonesia	0.321	0.268	0.830	0.563	1	0.034	0.204	16.95%	1	3
Madagascar	2.107	1.902	4.027	2.125	6	0.005	0.120	11.49%	4	8
Mauritius	0.929	0.482	4.420	3.938	3	0.003	0.637	63.40%	2	8
Mayotte	0.250	0.063	0.688	0.625	1	0.003	0.155	15.11%	1	2
Mozambique	0.893	0.661	2.473	1.813	3	0.007	0.240	23.34%	2	5
Namibia	0.000	0.000	0.438	0.438	0	0.000	0.346	34.58%	0	2
Réunion	0.429	0.170	3.170	3.000	3	0.000	0.390	38.94%	1	6
Seychelles	0.107	0.089	0.777	0.688	1	0.006	0.185	17.92%	1	2
South Africa	0.036	0.009	0.946	0.938	1	0.000	0.245	24.47%	0	3
Tanzania	0.036	0.009	0.759	0.750	1	0.000	0.179	17.86%	0	2
Timor-Leste	0.143	0.098	0.473	0.375	1	0.007	0.087	7.98%	1	2
Zimbabwe	0.107	0.089	1.152	1.063	1	0.006	0.318	31.22%	1	3

Table A 7: South Pacific - Prior distribution, posterior distribution, and posterior predictive distribution from observed data and counterfactual data

Country	Mean prior μ_{hist}	Mean posterior observed $E[\lambda|y_{obs}]$	Mean posterior worst-case CF $E[\lambda|y_{CFmax}]$	Difference posterior $E[\lambda|y_{CFmax}] - E[\lambda|y_{obs}]$	Max. observed y 80-07 (ymax)	$p(\widetilde{y} \geq ymax| y_{obs})$	$p(\widetilde{y} \geq ymax| y_{CFmax})$	$\Delta p(\widetilde{y} \geq ymax)$ [%]	$E[\widetilde{y}_{95p}|y_{obs}]$	$E[\widetilde{y}_{95p}|y_{CFmax}]$
American Samoa	0.321	0.393	1.455	1.063	2	0.010	0.184	17.38%	2	4
Australia	2.321	2.268	5.705	3.438	5	0.033	0.502	46.97%	5	10
Cook Islands	0.393	0.286	3.036	2.750	2	0.005	0.578	57.30%	1	6
Fiji	0.821	0.955	3.143	2.188	3	0.020	0.384	36.38%	3	6
French Polynesia	0.643	0.286	3.598	3.313	6	0.000	0.079	7.91%	1	7
Indonesia	0.036	0.134	1.196	1.063	1	0.011	0.334	32.27%	1	3
New Caledonia	1.071	0.830	4.018	3.188	4	0.003	0.374	37.19%	3	8
New Zealand	0.179	0.107	3.670	3.563	3	0.000	0.495	49.52%	1	7
Niue	0.036	0.009	1.946	1.938	1	0.000	0.571	57.10%	0	5
Papua New Guinea	0.536	0.196	1.509	1.313	3	0.000	0.072	7.22%	1	4
Solomon Islands	1.214	0.679	1.866	1.188	5	0.000	0.015	1.49%	2	4
Timor-Leste	0.000	0.063	0.750	0.688	0	0.059	0.517	45.81%	1	2
Tonga	0.429	0.607	2.795	2.188	2	0.028	0.524	49.55%	2	6
Tuvalu	0.107	0.027	0.152	0.125	1	0.001	0.014	1.25%	0	1
Vanuatu	1.429	1.045	3.795	2.750	4	0.006	0.332	32.62%	3	7

Countries – Figures Posterior Predictive distribution
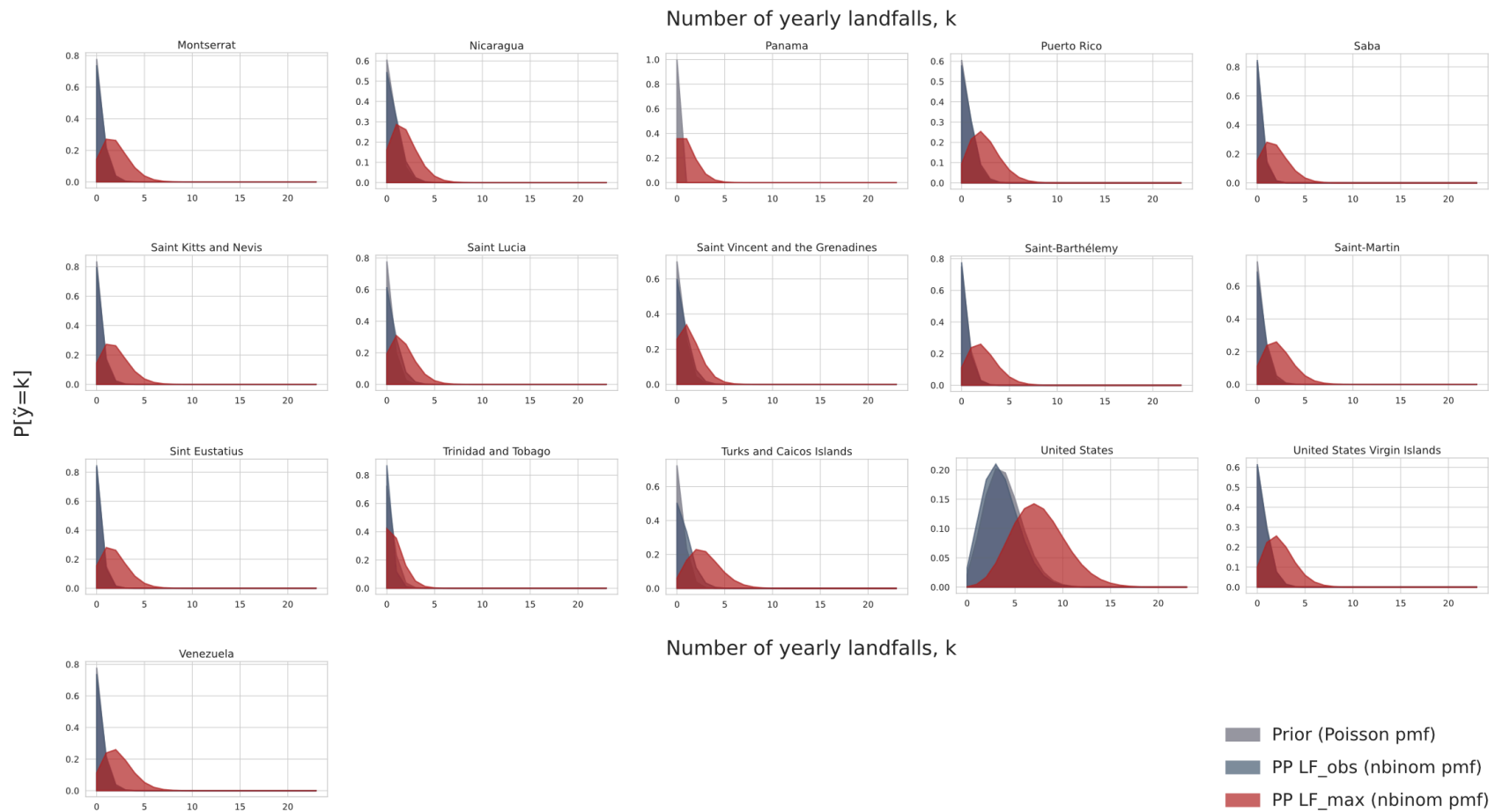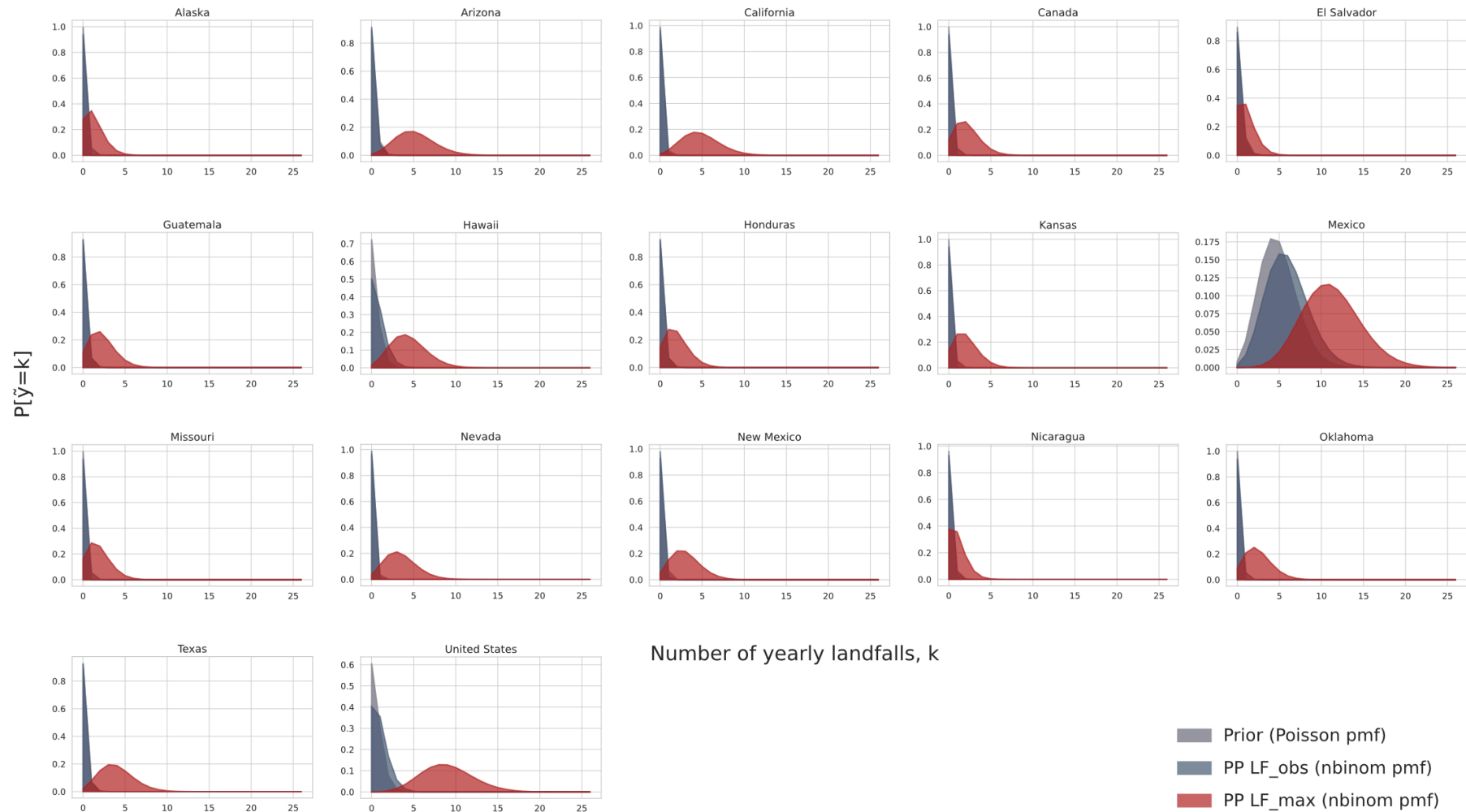

P[ỹ=k]

Number of yearly landfalls, k
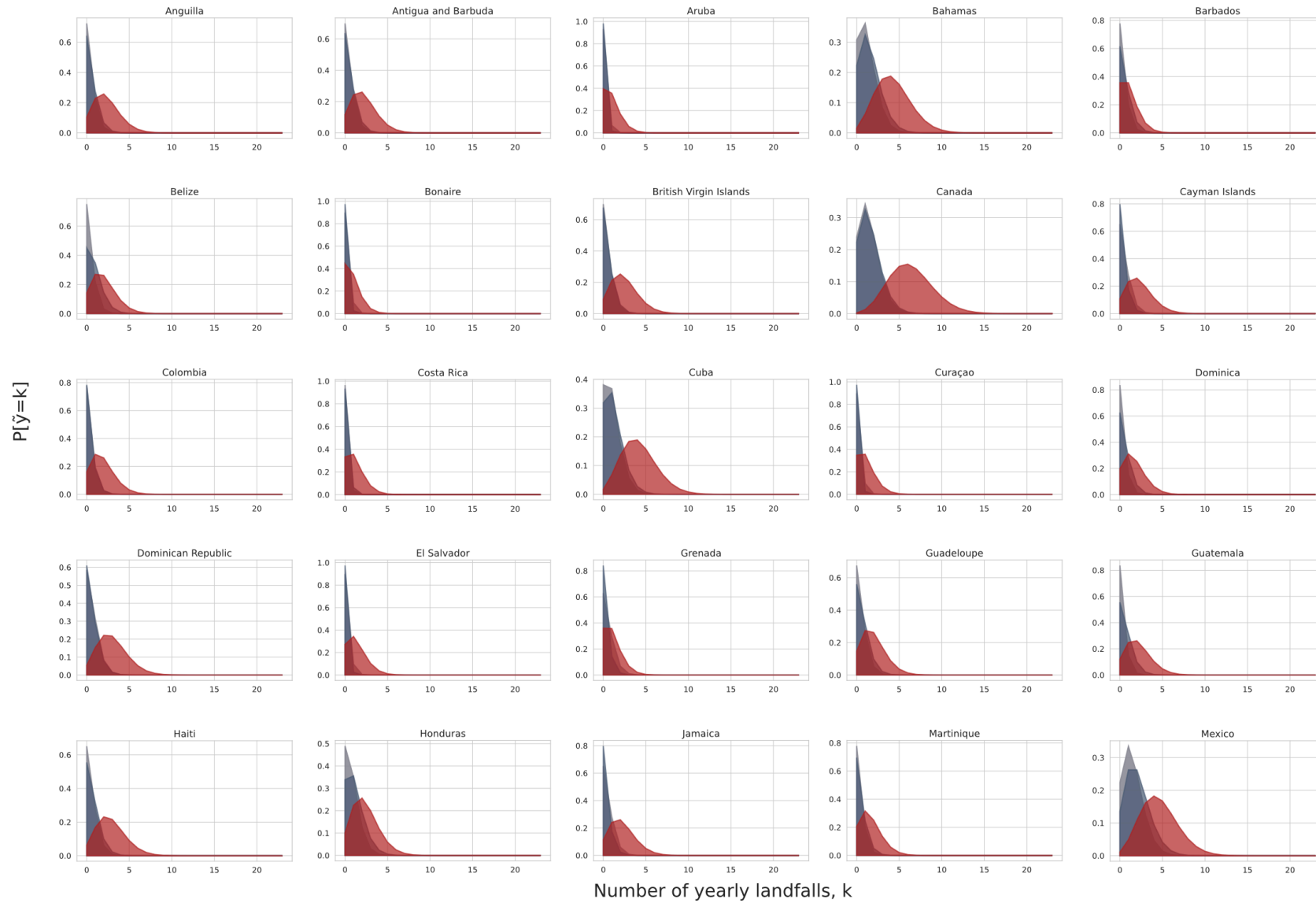
Figure A 1: North Atlantic - Posterior predictive distributions based upon different histories for 2008-2019. In blue is posterior predictive distribution based on observed data. In different shades of red are the individual posteriors based on counterfactual scenarios from the best-case scenario (left, light red) to the worst-case scenario (right, dark red). Grey shading is prior distribution as a reference. All distributions are based on the prior using historic data from 1980-2019 and forecast data from the combined dataset KWBC+ECMWF.
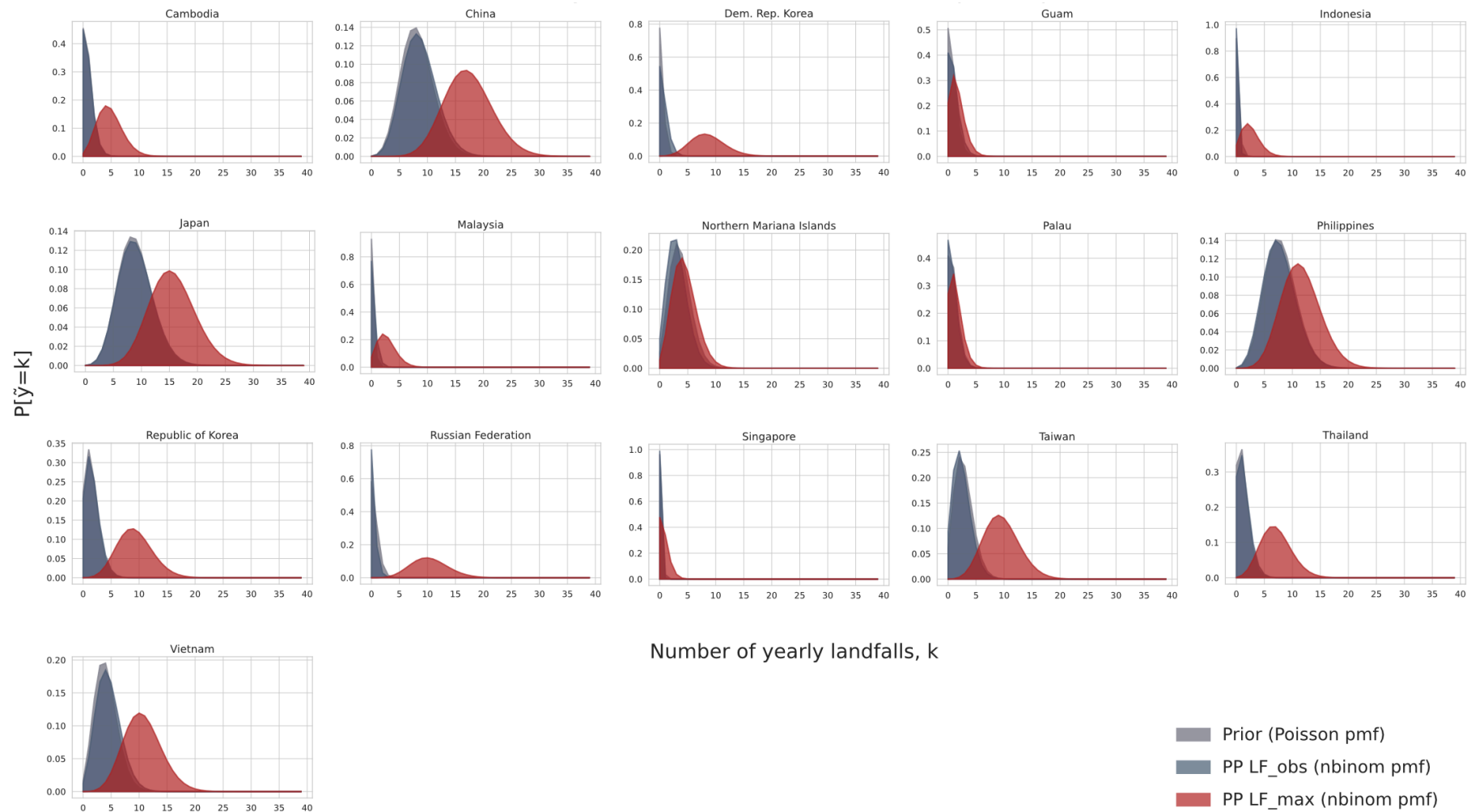
Figure A 2: East Pacific - Posterior predictive distributions based upon different histories for 2008-2019. In blue is posterior predictive distribution based on observed data. In different shades of red are the individual posteriors based on counterfactual scenarios from the best-case scenario (left, light red) to the worst-case scenario (right, dark red). Grey shading is prior distribution as a reference. All distributions are based on the prior using historic data from 1980-2019 and forecast data from the combined dataset KWBC+ECMWF.
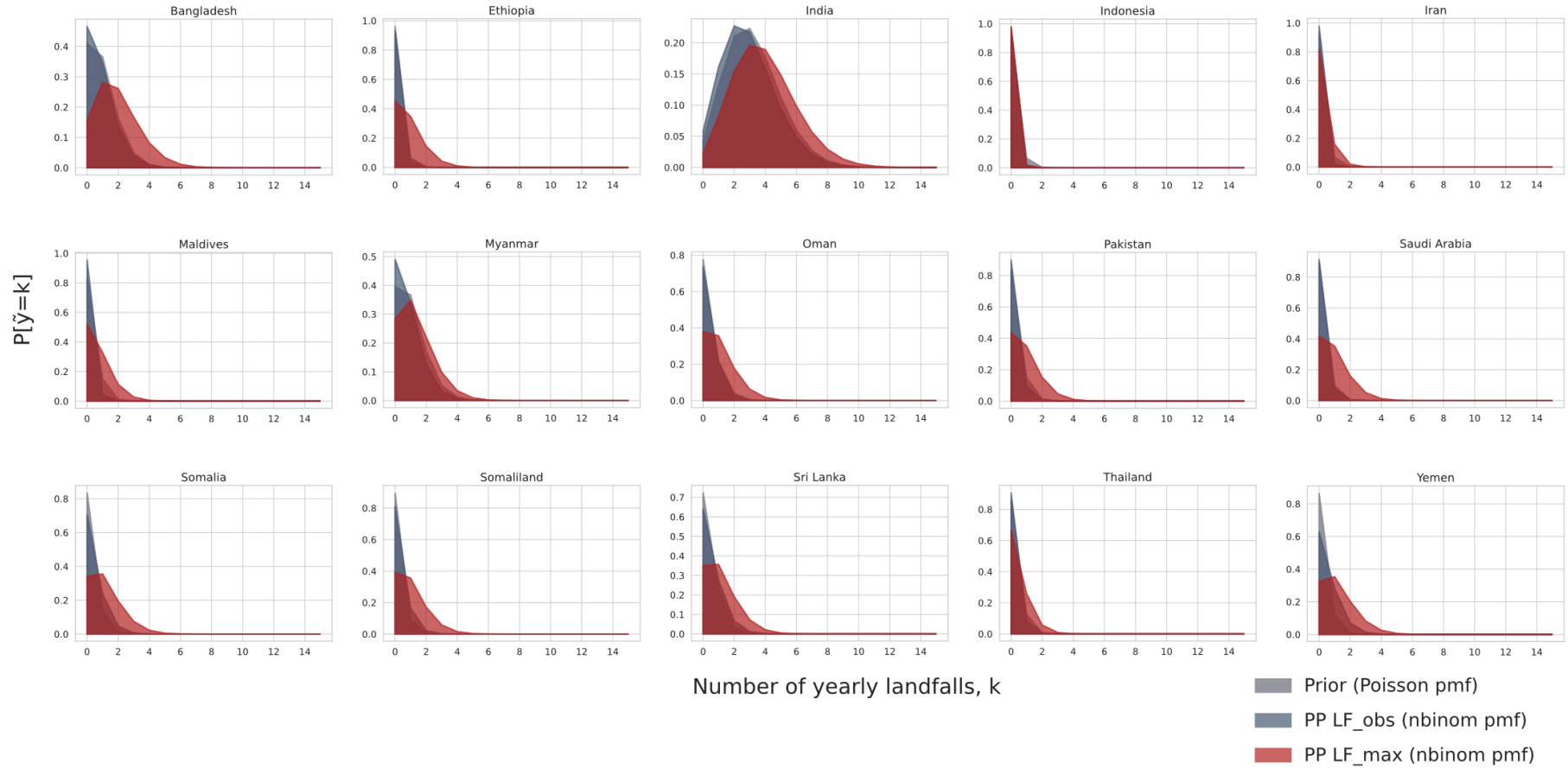
Figure A 3: West Pacific - Posterior predictive distributions based upon different histories for 2008-2019. In blue is posterior predictive distribution based on observed data. In different shades of red are the individual posteriors based on counterfactual scenarios from the best-case scenario (left, light red) to the worst-case scenario (right, dark red). Grey shading is prior distribution as a reference. All distributions are based on the prior using historic data from 1980-2019 and forecast data from the combined dataset KWBC+ECMWF.
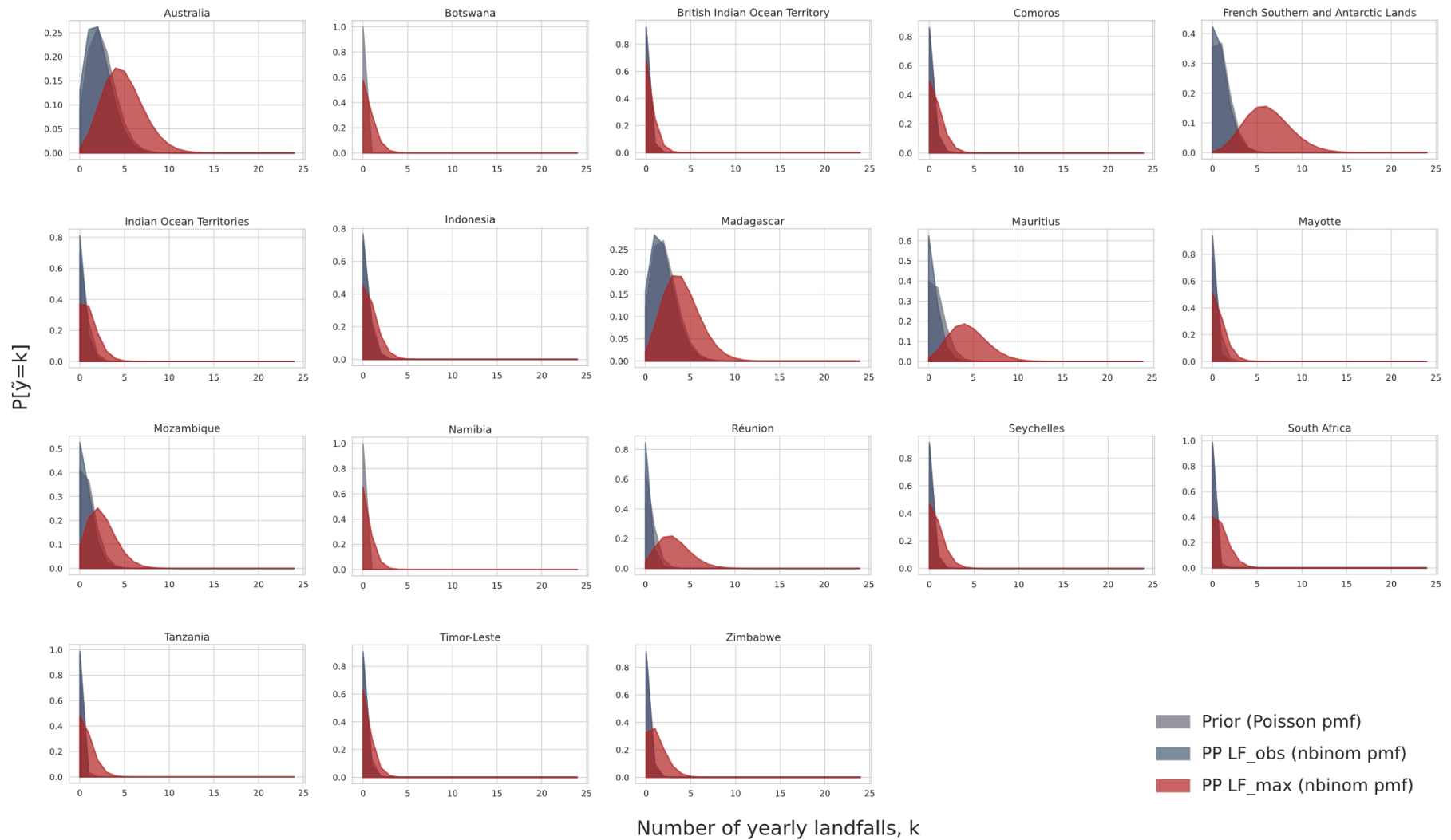
Figure A 4: North Indian Ocean - Posterior predictive distributions based upon different histories for 2008-2019. In blue is posterior predictive distribution based on observed data. In different shades of red are the individual posteriors based on counterfactual scenarios from the best-case scenario (left, light red) to the worst-case scenario (right, dark red). Grey shading is prior distribution as a reference. All distributions are based on the prior using historic data from 1980-2019 and forecast data from the combined dataset KWBC+ECMWF.

Figure A 5: South Indian Ocean - Posterior predictive distributions based upon different histories for 2008-2019. In blue is posterior predictive distribution based on observed data. In different shades of red are the individual posteriors based on counterfactual scenarios from the best-case scenario (left, light red) to the worst-case scenario (right, dark red). Grey shading is prior distribution as a reference. All distributions are based on the prior using historic data from 1980-2019 and forecast data from the combined dataset KWBC+ECMWF.
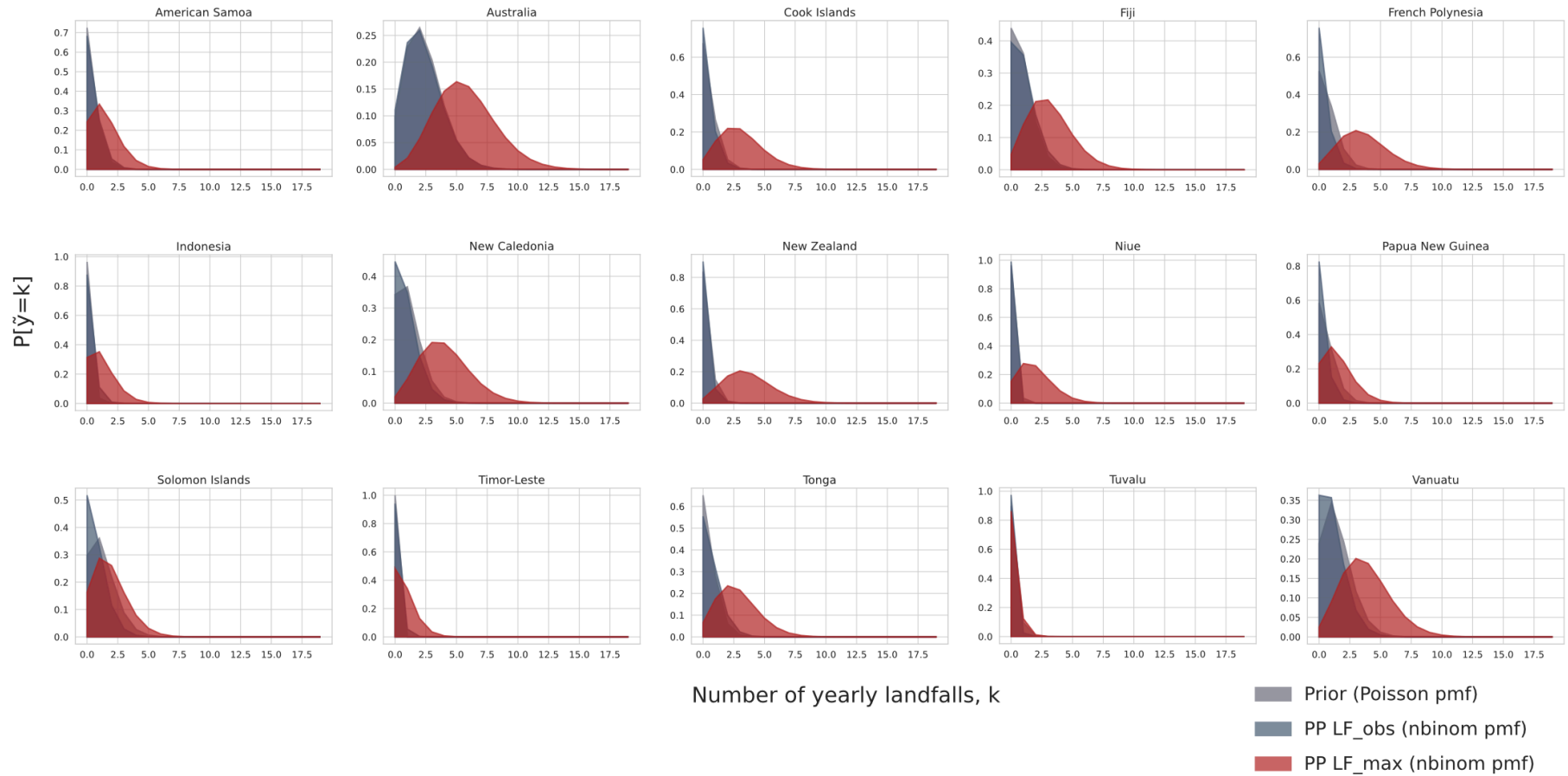
Figure A 6: South Pacific - Posterior predictive distributions based upon different histories for 2008-2019. In blue is posterior predictive distribution based on observed data. In different shades of red are the individual posteriors based on counterfactual scenarios from the best-case scenario (left, light red) to the worst-case scenario (right, dark red). Grey shading is prior distribution as a reference. All distributions are based on the prior using historic data from 1980-2019 and forecast data from the combined dataset KWBC+ECMWF.

Declaration of consent

on the basis of Article 30 of the RSL Phil.-nat. 18

Name/First Name: Baumann Tamara Alessandra

Registration Number: 14-122-444

Study program: Climate Sciences

Bachelor ☐   Master ☑   Dissertation ☐

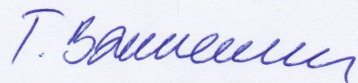Title of the thesis: Counterfactual Risk Analysis of Tropical Cyclones

Supervisor: Prof. Dr. Olivia Romppainen-Martius

I declare herewith that this thesis is my own work and that I have not used any sources other than those stated. I have indicated the adoption of quotations as well as thoughts taken from other authors as such in the thesis. I am aware that the Senate pursuant to Article 36 paragraph 1 litera r of the University Act of 5 September, 1996 is authorized to revoke the title awarded on the basis of this thesis.

For the purposes of evaluation and verification of compliance with the declaration of originality and the regulations governing plagiarism, I hereby grant the University of Bern the right to process my personal data and to perform the acts of use this requires, in particular, to reproduce the written thesis and to store it permanently in a database, and to use said database, or to make said database available, to enable comparison with future theses submitted by others.

Place/Date

Bern, 4.2.22

T. Baumann

Signature