RESEARCH ARTICLE

A data science approach to climate change risk assessment applied to pluvial flood occurrences for the United States and Canada

Mathilde Bourget[1,2], Mathieu Boudreault[1] , David A. Carozza[1], Jérémie Boudreault[3,4] and Sébastien Raymond[3,4]

[1]Department of Mathematics, Université du Québec à Montréal, Montréal, QC, Canada
[2]Collège Jean-de-Brébeuf, Montréal, QC, Canada
[3]Climatic Hazards and Advanced Risk Modelling, Co-operators General Insurance Company, Québec, QC, Canada
[4]Centre Eau Terre Environnement, Institut national de la recherche scientifique, Québec, QC, Canada
Corresponding author: Mathieu Boudreault; Email: boudreault.mathieu@uqam.ca

Abstract

There is mounting pressure on (re)insurers to quantify the impacts of climate change, notably on the frequency and severity of claims due to weather events such as flooding. This is however a very challenging task for (re)insurers as it requires modeling at the scale of a portfolio and at a high enough spatial resolution to incorporate local climate change effects. In this paper, we introduce a data science approach to climate change risk assessment of pluvial flooding for insurance portfolios over Canada and the United States (US). The underlying flood occurrence model quantifies the financial impacts of short-term (12–48 h) precipitation dynamics over the present (2010–2030) and future climate (2040–2060) by leveraging statistical/machine learning and regional climate models. The flood occurrence model is designed for applications that do not require street-level precision as is often the case for scenario and trend analyses. It is applied at the full scale of Canada and the US over 10–25 km grids. Our analyses show that climate change and urbanization will typically increase losses over Canada and the US, while impacts are strongly heterogeneous from one state or province to another, or even within a territory. Portfolio applications highlight the importance for a (re)insurer to differentiate between future changes in hazard and exposure, as the latter may magnify or attenuate the impacts of climate change on losses.

1. Introduction

There is growing pressure on the financial services industry (insurers, reinsurers, and banks) to factor in climate extremes and climate change in their business decisions. This is because corporations globally are integrating ESG (environmental, social, and corporate governance) principles and will soon become subject to a new regulatory environment, thanks in large part to the work of the Task Force on Climate-related Financial Disclosures (TCFD) (Financial Stability Board, 2017). Regulators are gradually requiring corporations to report the sensitivity of their profitability to various climate scenarios, whereas banks and (re)insurers must do similarly by stress-testing their stability as well (e.g., Bank of England, 2019; OSFI, 2023).

An important component of such reporting and stress-testing for the property (re)insurance industry is climate change risk assessment (CCRA), that is, quantifying the financial impact of climate change on the frequency and intensity of claims due to, for example, flooding. CCRA is very challenging to actuaries, catastrophe modelers, and financial risk managers because it requires modeling and understanding

of the financial impacts of climate change at the scale of a (re)insurance portfolio (country-wide or globally) and at a spatial resolution that allows the integration of local climate change effects (regional, city).

Flooding is the most significant natural hazard in the United States (US) and Canada (FEMA, 2017; Canada, 2022b), and CCRA of flooding is especially challenging. Namely, large-scale high-resolution flood modeling is (1) extremely costly from both a computational and financial standpoint; (2) they require inputs that are difficult to acquire for large countries (high-resolution terrain and bathymetry data) or are nonexistent for many cities (sewer system configuration, exact location of inlets and outlets) and ultimately; and (3) they lack the flexibility required by, for example, actuaries and economists to analyze many customized scenarios (for a review, see Carozza and Boudreault, 2021 and references therein). Computing the impacts of climate change on flooding adds another layer of difficulty, typically requiring future projections of precipitation and temperature from climate models. However, running climate models is very time-consuming, which in turn limits their spatial resolution and the number of runs available. Hence, for financial risk management applications that do not necessarily require accurate street-level data, such as analyzing scenarios and trends for (re)insurance portfolios over Canada and/or the US, the resolution provided by regional climate models is appropriate.

In this paper, we introduce a data science approach to CCRA of pluvial flooding for insurance portfolios over Canada and the US. Pluvial flooding, that is, "heavy rainfall-related flooding that is independent of an overflowing body of water" (ICLR, 2021) is different from the processes that lead to fluvial flooding or the overflow of a river or watercourse. The underlying flood model is therefore focused on quantifying the financial impacts of short-term (12–48 h) precipitation dynamics over the present and future climate (until 2060) using a methodology that leverages statistical/machine learning and climate models. This is done through a top-down modeling chain integrating climate model outputs at its core. Few papers in actuarial science have integrated climate models for insurance applications. For example, Boudreault *et al*. (2020) used a top-down modeling approach and a chain of climate, hydrological, and hydraulic models to represent fluvial flood risk over a small city in Canada. Jin and Erhardt (2020) used climate model outputs to price temperature index-based insurance products in California. Here, the analysis is performed for Canada and the US over a 10- or 25-km grid depending on the application, keeping an appropriate balance between computational speed and the ability to distinguish regional discrepancies.

To meet such goal, we trained statistical and machine learning methods on historical pluvial flood occurrences in the US and validated their predictive skill over the US (test set) and Canada (validation set). For a review of applications of statistical/machine learning and artificial intelligence in actuarial science, see, for example, Yeo *et al*. (2019), Blier-Wong *et al*. (2020), and Richman (2021a,b) and in flood prediction, see, for example, Mosavi *et al*. (2018). We then integrated output from a regional climate model to calculate future flood probabilities for every month and grid cell until 2060. Finally, we show various portfolio applications where we analyze the impacts of changes in hazard and exposure on portfolio losses over the present (2010–2030) and future (2040–2060). The paper provides a methodology for CCRA that is applicable to various risks, which is achieved by connecting statistical and climate models to solve a problem that is increasingly important for both actuarial science and actuaries.

Overall, we find that generalized additive models (GAMs) have a solid predictive power in- and out-of-sample to explain pluvial flood episodes compared to linear models and ensemble tree-based methods. We find different levels of pluvial flood risk in most urban areas of Canada and the US. Furthermore, we do not recommend using tree-based methods for projecting the impacts of future precipitation and urbanization patterns due to their inability to extrapolate beyond the original training set. Our work highlights a wide heterogeneity of climate change impacts across states and provinces that become significant when analyzing insurance portfolios, highlighting the importance of climate-informed financial risk management. We also emphasize the importance of differentiating changes in hazard and exposure since they both interact to attenuate or magnify the financial impacts of climate change.

The paper is structured as follows. Section 2 describes physical risk assessment (PRA), CCRA, and how it is applied in the context of this paper. Section 3 then details the datasets, statistical and machine learning methods used to build a variety of pluvial flood models, and evaluates their predictive power.

We then show in Section 4 future projections of pluvial flood risk over Canada and the US, and for selected cities in both countries. We present a portfolio application in Section 5 that highlights how regional discrepancies and portfolio composition may affect aggregate losses. Section 6 then concludes with a broad discussion of the paper's findings and the models' limitations. Finally, the Supplementary Material (SM) completes the core analyses of the paper by providing additional results and validations.

2. Physical risk assessment

PRA in the context of this paper is the qualitative and quantitative analysis of the impacts of climatic events such as floods, tropical cyclones, and wildfires. For a property and casualty insurance organization (public or private), PRA requires an understanding of the frequency and intensity of these events, without or with climate change considerations, and their impact on the claims dynamics. While PRA and CCRA are related, they mostly differ on the time horizon of the analysis: PRA centers on immediate physical risks to properties, whereas CCRA focuses on potential future risks tied to climate change. In both cases, PRA and CCRA are typically done through a decomposition of risk into its main components of hazard, vulnerability, and exposure (Mitchell-Wallace *et al.,* 2017).

2.1 Top-down catastrophe modeling

According to the United Nations Office for Disaster Risk Reduction (UNDRR) Sendai Framework Terminology on Disaster Risk Reduction (UNDRR, 2017), hazard, vulnerability, and exposure are defined as (IPCC, 2021a uses a similar terminology):

- "Hazard: a process, phenomenon or human activity that may cause loss of life, injury or other health impacts, property damage, social and economic disruption or environmental degradation;
- Vulnerability: the conditions determined by physical, social, economic and environmental factors or processes which increase the susceptibility of an individual, a community, assets or systems to the impacts of hazards;
- Exposure: the situation of people, infrastructure, housing, production capacities and other tangible human assets located in hazard-prone areas."

Risk is at the intersection of hazard, vulnerability, and exposure. This is well illustrated in Figure 1 of UN (2023) or Figure TS.4 of IPCC (2021b). For example, a property is exposed to flooding if it is located in an area of flood hazard, whereas it is vulnerable to flooding if there are possible entries where water can enter into a house (basement windows or doors). Catastrophe modeling is based on such decomposition of risk and aims to model each of these three components, providing in the end what is known as the ground-up loss; that is, losses before the application of any insurance or reinsurance. This is illustrated at the bottom of Figure 1.

The hazard component represents the frequency, intensity, duration, and footprint of an event. The exposure includes the geographical location of the property, its size (e.g., square footage, number of floors) and value (e.g., market value, reconstruction costs). Vulnerability represents the characteristics of a house that magnifies or attenuates the impacts of the hazard. In the context of flooding, this includes for example whether the basement is finished/unfinished or there is a crawlspace, the first floor elevation, and the height of basement windows, etc. Damage curves typically link the intensity of an hazard with the vulnerability of a home to yield dollars of losses or percentage of damage.

2.2 Climate models

Hazard modeling of climatic events such as flooding or tropical cyclones is founded on an understanding of the interactions between the climate and, for example, the frequency and intensity of a climatic hazard
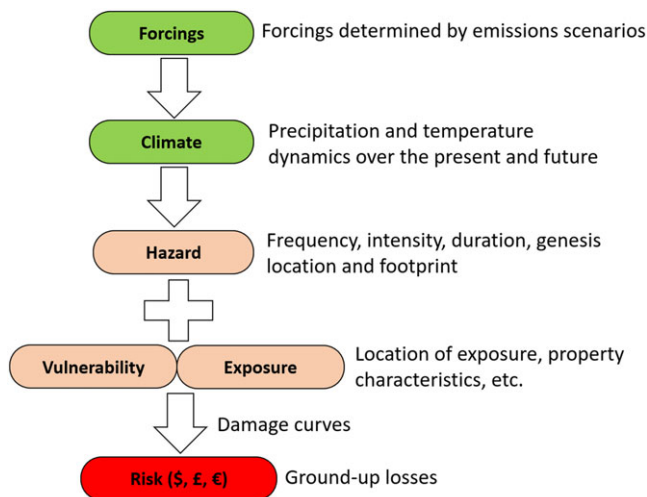
Figure 1. *Top-down catastrophe modeling approach with climate on top.*

(top of Figure 1). CCRA adds another layer of modeling as we need to relate greenhouse gas (GHG) emissions and concentration to impacts on key climate variables such as temperature and precipitation. A natural approach is therefore the integration of climate models (general circulation models and regional climate models) into a PRA.

General circulation models (GCMs) are numerical models that simulate the evolution and interactions between the components of the climate system (atmosphere, land, ocean, ice, etc.) using physical equations and empirical relationships (Chen *et al.,* 2021). They are at the core of climate and climate change studies and are thus widely used to study global temperature and precipitation (among other variables) patterns over the present and future (Chen *et al.,* 2021). In many respects, regional climate models (RCMs) are similar to GCMs yet they model atmospheric phenomena at continental and regional scales allowing simulations at higher spatial and temporal resolutions, which can resolve smaller-scale physical processes (Chen *et al.,* 2021). Climate models are computationally intensive and typically run on large supercomputing clusters.

Climate models are forced with GHG emissions scenarios that are designed to capture the impacts of future socioeconomic growth and energy alternatives. Those emissions scenarios are in turn converted into radiative forcings (see top of Figure 1). Changes in the radiative forcing represent the extra warming in the atmosphere due to GHGs and are measured in watts per square meter. For example, the IPCC AR6 SSP2-4.5 scenario, which represents a "Middle of the Road" future (see Fricko *et al.,* 2017 for the detailed storyline), corresponds to an extra energy flux of 4.5 Wm^{-2} to the atmosphere by 2100. Over available climate models, this emissions scenario typically leads to an approximate global warming of +2.5 degrees in 2100 compared to preindustrial.

Typical outputs of climate models include (surface) temperature, (liquid, snow, and convective) precipitation, (surface) relative/specific humidity, eastward and northward (surface) winds, (surface) air pressure, etc. Outputs typically are stored as grids that may take up to terabytes and petabytes of storage depending on the vertical (in the atmosphere), horizontal (over the surface), and temporal (hourly, daily) resolution of the data and the number of variables stored.

Integrating climate models into hazard modeling should also take into consideration the biases in the outputs that can affect the results at the end of the modeling chain. This is typically approached using what is called preprocessing or postprocessing. Preprocessing implies bias correction prior to its application in a hazard module. This is done by comparing, for example, simulated precipitation outputs with past observations and applying an appropriate correction. Postprocessing implies comparing

a hazard feature simulated from climate model outputs (e.g., hazard frequency) with what was observed in the past. If preprocessing does not succeed in eliminating all biases in the hazard component, then postprocessing could also be applied.

Reporting and regulatory requirements are often based on assessing the overall impacts to an organization of a given temperature increase (e.g., an increase of 2 degrees compared to preindustrial temperature). Temperature increase and global warming are endogenous in a climate model and result from radiative forcings. Integrating climate models into PRA and CCRA therefore requires using available runs of climate models along with corresponding climate scenarios.

2.3 Proposed approach

We take a data science approach to CCRA for pluvial flooding in an insurance portfolio. The core of the work lies in the hazard modeling of pluvial flooding occurrence (Y) as a function of a set of atmospheric and socioeconomic variables (X). We first fit and validate the $Y|X$ relationship using statistical and machine learning approaches for binary responses using past observed data for X and Y. This is the fitting and validation step. Climate (change) risk assessment is then performed by computing flood occurrence probabilities using outputs from climate models for X over different time intervals. This is the projection and simulation step. When computing predictions, one may hold socioeconomic variables fixed to isolate the effects of climate change from socioeconomic growth, or one may use projections of the socioeconomic variables.

This study is not meant to provide a detailed account of the impacts of pluvial flooding at the street level. There is inevitably a trade-off between the financial and/or computing resources necessary and the resolution of the information needed. Here, we focus on the large-scale impacts of climate change on pluvial flooding in portfolios covering Canada or the US. As such, we do not explicitly model vulnerability, yet we proxy exposure as the number of people living or insured in each grid cell. More details about portfolio modeling are provided in Section 5.

3. Occurrence models

This section describes the pluvial flood occurrence models analyzed throughout the paper. We begin by outlining the datasets (Section 3.1), the statistical and machine learning methods (Section 3.2). We then explain how they have been applied to our study (Section 3.3) and complete the section by assessing the predictive capability of the models in the US and Canada (Sections 3.4.1 and 3.4.2).

3.1 Data

This section characterizes the historical flood occurrence data (Section 3.1.1) used as the response variable in the statistical and machine learning models. Then we examine the predictors derived from atmospheric and socioeconomic variables (Sections 3.1.2 and 3.1.3).

3.1.1 Flood occurrence

Historical flood occurrence is derived from the Storm Events Database (SED) from the National Oceanic and Atmospheric Administration (NOAA) (NOAA, 2021). The dataset contains significant weather events from 1951 onward in the US. No similar dataset is available for Canada, and this is discussed in Section 3.4.2. Information used to compile the SED comes from multiple sources, that is, 911 call centers, media, and local authorities such as law enforcement. For each event, there are many variables available, such as the date and time of the beginning and end of the event, the location of the event, and the type of event. For flood events, the database includes the cause of the flooding. For the purpose

of this research, we focused on flooding events induced by heavy rain between 2007 and 2020 because latitude/longitude location data was not available prior to 2007.

We then converted the storm events location data into monthly grids over the US such that there is at most one event per grid cell (and per month). We chose a monthly observation frequency to capture seasonality while keeping the overall size of the dataset manageable. We fixed the grid cell size to $0.1° \times 0.1°$ to match the precipitation data that was employed (Section 3.1.2). This spatial grid is approximately equivalent to 10 km $\times$ 10 km over the area of study (but as we approach the North Pole, 1 degree of longitude is much smaller and therefore, the grid cell area expressed in km^2 decreases as well). Therefore, historical flood occurrence data (and covariates as well) is represented over 168 grids (1 per month over 14 years) of 143,922 cells each.

3.1.2 Atmospheric variables

Precipitation is obviously a key driver of pluvial flooding, and data is extracted from the Multi-Source Weighted-Ensemble Precipitation (MSWEP), a comprehensive dataset that combines rain gauges, satellites, and reanalyses from various sources (Beck *et al.,* 2019). The 3-hourly data is available globally at a spatial resolution of $0.1° \times 0.1°$ from 1979 onward, but we exployed data from 2007 to 2020 to match the flood occurrence data. We constructed precipitation covariates by computing the monthly maximum of 6-, 9-, 12-, 24-, and 48-hourly precipitation.

Temperature is an important driver of evapotranspiration that also captures seasonal features of flooding. Temperature data comes from the CPC Global Daily Temperature data from the NOAA (National Oceanic and Atmospheric Administration et al., 2021). The dataset contains global gridded daily minimum and maximum temperatures at a resolution of $0.5° \times 0.5°$ from 1979 to the present. Because temperature resolution is lower than precipitation, we downscaled the data to $0.1° \times 0.1°$ assuming constant average daily temperature within each block of 5×5 grid cells. Given that spatial variations of temperature are typically much smaller than precipitation, this is a reasonable assumption. We constructed temperature covariates by recording the monthly average of daily minimum and maximum temperature at each grid cell.

In Northern or alpine climates, rapid snowmelt and rain-on-snow events caused by heavy rain (and rapid increases in temperature) are drivers of flooding. We therefore include snow cover data in the analysis from the Canadian Centre for Climate Services available over North America that is available over a 24 km $\times$ 24 km grid (Ross and Bruce, 2010). Snow cover data has been reprojected with bilinear interpolation to the same $0.1° \times 0.1°$ grid that we use for the analysis. Every month, we recorded the daily maximum snow cover.

There are many common climate types in the US and Canada, and we aimed to distinguish pluvial flooding dynamics based on such climates using the Köppen-Geiger (KG) climate classification. There are 30 climates spread over 5 main climate groups: tropical, dry, temperate, continental, and polar. KG climate classification is available as a static variable on a $0.1° \times 0.1°$ global grid (Peel *et al.,* 2007). Climate classification is not meant to be dynamic but rather to distinguish geographical areas based on weather patterns. Therefore, we assume it remains constant during the study period.

3.1.3 Socioeconomic variables

Land use is an important driver of flooding that determines how rainfall runs off the surface. Urbanization has led to an increase in flooding in the past (Feng *et al.,* 2021) by limiting infiltration and increasing surface runoff. We apply the land use data from the Commission for Environmental Cooperation (CEC, 2015), which was derived from data from the Landsat satellites for the year 2015 over a 30 m $\times$ 30 m grid.

Land use is divided into 19 classes. For each $0.1° \times 0.1°$ grid cell, there are over 100,000 observations of land use. We have therefore computed the proportion of each land use type for each of the 19 classes assuming it did not change significantly over the 14 years covered by this study. For parsimony, we

grouped 11 types of land use together, leaving us with 8 categories overall: forest, scrub, grassland, wetland, cropland, dry land, urban area, and water.

As floods are only reported when there is population and because we lack appropriate projections of land use for the future, we have also included population data into the analysis. We used the US Census Grid population data available for the years 2000 and 2010 from the Socioeconomic Data and Applications Center (SEDAC) hosted at Columbia University (Seirup and Yetman, 2006; CIESIN, 2017). The 30 arc-second (about 1 km) grid of the US Census was then aggregated to a $0.1° × 0.1°$ grid. A linear interpolation was used to deduce population for 2007 to 2009, and it was assumed fixed as of 2010 between 2011 and 2020 because 2020 population data was not available at the time of study. Note that gridded population data is approximately equivalent to population density since the grid cell size remains constant at $0.1° × 0.1°$.

3.2 Methods

Flood occurrence is a classification problem and as such, we have used the generalized linear model (GLM) (namely, the logistic regression), the GAM, and random forests (RF). The GLM extends multiple regression models for, for example, binary or count responses. It does so by modeling a transformation of the expected response as a linear function of the predictors. The GAM expands on the GLM by allowing nonlinear transformations of the predictors. Finally, RF combines decision trees by sampling observations and predictors to derive empirical relationships. We focused on the latter three methods because the resulting models are flexible and interpretable while allowing for nonlinearities (GAM and RF) and interactions (RF). More details about GLM, GAM, and RF methods can be found in Chapters 4, 7, and 8 of James *et al*. (2021) and in the SM Section 1.1.

The historical occurrence data described in Section 3.1.1 has more than 2.4 million observations, of which 0.27% are ones (pluvial flood observed in a given month and grid cell) and the rest are zeroes (no pluvial flood observed in a given month and grid cell). As such, the dataset is imbalanced and we focus on avoiding problems related to the overestimation of the probability of no flood (false negatives).

There are few solutions to deal with data imbalance (Ganganwar, 2012). It is possible to oversample ones or undersample zeroes. Given the size of the dataset, undersampling zeroes was a more appropriate approach than oversampling ones, since it reduces the database size and accelerates computations. In other words, we randomly (over months and grid cells) eliminated zeroes from the dataset so that the resulting proportion of zeroes was either 90% or 50%. We tested two proportions to determine whether the outcomes are sensitive to this choice. Whenever undersampling was used, predicted flood occurrence probabilities were adjusted following Saerens *et al*. (2002) to match observed probabilities.

3.3 Models

The response variable is flood occurrence in the US, measured over grid cells and months. We assume that given a set of covariates, flood occurrence is independent (over grid cells of $0.1° × 0.1°$ and months) and as such the problem becomes a typical classification problem (a discussion on this assumption is included in Section 6). The set of covariates are (as described in Section 3.1) five precipitation variables (6-hourly, 9-hourly, 12-hourly, 24-hourly, and 48-hourly), temperature, snow cover, climate classification, eight land use proportions, and population density.

To eliminate the adverse effects of multicollinearity, we constructed a smaller set of covariates. We found high correlation in the precipitation variables (by construction, as one precipitation variable was often included in another) and only kept the 24-h precipitation since higher frequency precipitation is not available in the climate model projections described in Section 4.1. Moreover, we found high correlation between monthly temperature and snow cover, which led us to exclude snow cover since temperature is readily available in climate models whereas snow cover is not. We also found high correlation between the proportion of urban extent (one of the land use covariates) and population density and decided to

keep the latter since it is readily available in future population projections. We also combined forest and grassland proportions into a single land use category for similar reasons. Overall, the smaller set of covariates comprises of 24-h precipitation, average daily temperature, climate classification, five proportions of land use (combining forest and grassland, dropping water, and urban extent), and population density.

Given the imbalanced dataset problem and correlation between many of the covariates (especially between the precipitation variables), we analyze each of the following using the GLM, GAM, and RF:

- All covariates, no undersampling;
- All covariates, undersampling with 90% zeroes;
- All covariates, undersampling with 50% zeroes;
- Smaller set of covariates, undersampling with 90% zeroes;
- Smaller set of covariates, undersampling with 90% zeroes, with logged population density.

Consequently, there are 15 models overall.

3.4 Validation

We first fit the 15 flood occurrence models (described in Section 3.3) over the US with 70% of data and assess the predictive capability using a test set made of the remaining 30% (Section 3.4.1). The test set has been generated by randomly selecting 30% of the 2.4 million observations, where one observation is either a 1 or 0 for each grid cell and month. In other words, the random sampling is achieved over both the temporal and spatial dimensions simultaneously. The training set is therefore made of the remaining observations. More details about the implementation of the GLM, GAM, and RF can be found in the SM.

A model fitted over the US (with Canadian KG climates available in the US) is then used to predict pluvial flooding in Canada. The model quality over Canada is investigated in two ways (Section 3.4.2). First, we used flood claims data from a major Canadian insurance company, which yielded a purely out-of-sample predictive analysis. Second, we perform a qualitative assessment of the model over major historical flood events in Canada.

3.4.1 United States

We can evaluate the predictive power of occurrence models by determining whether a model can predict an occurrence when one was observed, or vice versa, predict a non-occurrence when nothing occurred. This is done by computing true/false positive/negative rates. Popular approaches involve analyzing the receiver operating characteristic (ROC) and precision-recall (PR) curves. The ROC curve plots the true positive rate as a function of the false positive rate, whereas the PR curve focuses on the relationship between precision and recall, which are two different ways to express the true positive rate. There are however limitations about the ROC curve (Cook, 2007; Saito and Rehmsmeier, 2015; Muschelli, 2020) notably when the dataset is imbalanced, whereas the ROC is only a valid model selection criterion under auto-calibration (see Wüthrich, 2023 and references therein). We will therefore use the PR curve as an additional metric to assess models' performance. For more details about the ROC and PR metrics, the reader should refer to Chapter 4 of James *et al.* (2021).

To evaluate the predictive power of the 15 occurrence models, we computed the areas under the ROC and PR curves with the test set in the US. The results are shown in Table 1.

We find that models perform very well in the test set with an area under the ROC curve in the range of 0.89–0.93, with a slight advantage to RFs. As for the area under the PR curve, values range from 0.05 to 0.11, which is above the baseline for a non-informative model (0.005, computed with historical occurrences). The GAM model performs better than the GLM, while RF again shows the best predictive

Table 1. *Area under the ROC and PR curves with the test set over the US for all 15 models considered. Note that "u/s" stands for undersampling.*

Models	ROC curve			PR curve		
	GLM	GAM	RF	GLM	GAM	RF
All, no u/s	0.8890	0.9076	**0.9224**	0.0587	0.0762	**0.1091**
All, 90% u/s	0.8992	0.9084	**0.9254**	0.0587	0.0742	**0.0923**
All, 50% u/s	0.9017	0.9085	**0.9216**	0.0566	0.0701	**0.0766**
Smaller, 90% u/s	0.8894	0.8987	**0.9273**	0.0499	0.0653	**0.0957**
Smaller, 90% u/s, log pop	0.8969	0.9038	**0.9272**	0.0564	0.0676	**0.0957**

capability in the test set. Across all models considered, RF yields the largest area under both the ROC and PR curves.

We find that there is no clear advantage from a predictive standpoint to undersample zeroes (Table 1). When using all covariates, undersampling with a target of 50% or 90% zeroes provided a very similar area under the ROC curve for the GLM, GAM, and RF. The latter result is different though for the PR curve, where undersampling slightly worsens predictive capability for the GAM, and more significantly for the RF. That said, the more parsimonious models with undersampling still yielded strongly comparable performance to cases without undersampling. When computation times matter, undersampling the dataset and using less covariates therefore yield very similar predictive performance (in a shorter amount of time and with less computational cost).

We compare in Figure 2 the predicted probabilities from one RF model with historical flood probabilities. We see that pluvial flood occurrence is concentrated in urban areas and that the model does very well in characterizing the spatial characteristics of pluvial flood, which is important to distinguish where climate change might have a greater impact. In the case illustrated here, the RF appropriately captures areas of low risk (white on top vs. yellow in the bottom) and pluvial flood probabilities in urban areas are at very similar levels and locations in both panels.

The SM includes 15 plots, one per model (GLM, GAM, and RF) and one per set of covariates (5), plus the empirical probability. They show that adding the logged population density was important for GLM and GAM, since it appears that predictions were too sensitive to slight changes in population otherwise. We also added to the SM a summary of model outputs for the GLM and GAM trained with the smaller set of covariates, with undersampling and logged population density. For both methods, precipitation is by far the most significant predictor. With the GLM, the coefficient for precipitation is positive and with the GAM the relationship is increasing nonlinearly. Other statistically significant predictors include population density, the proportion of wetlands and temperature, among others.

3.4.2 Canada

We would like to perform a validation exercise of the 15 models but over Canada. There are however no formal datasets in Canada that record flood events (or other weather events) at a level of granularity that we can find in the NOAA SED (with latitude and longitude of location). The Canadian Disaster Database (Canada, 2022a), maintained by Public Safety Canada, has approximately the same level of information as the EM-DAT dataset (Guha-Sapir *et al.*, 2022). The Flood List website (Davies *et al.*, 2021) also provides information about flooding globally, but in all three cases, location information is much too vague to be able to formally validate the flood models.

We thus perform a quantitative and a qualitative validation of the models over Canada. The quantitative assessment was based on a sample (non-random and non-divulged in this article) of data from Co-operators, a Canadian insurance company part of the top six P&C insurers in Canada. Data specific to clients were not used for this analysis, and only aggregate information about a flood event
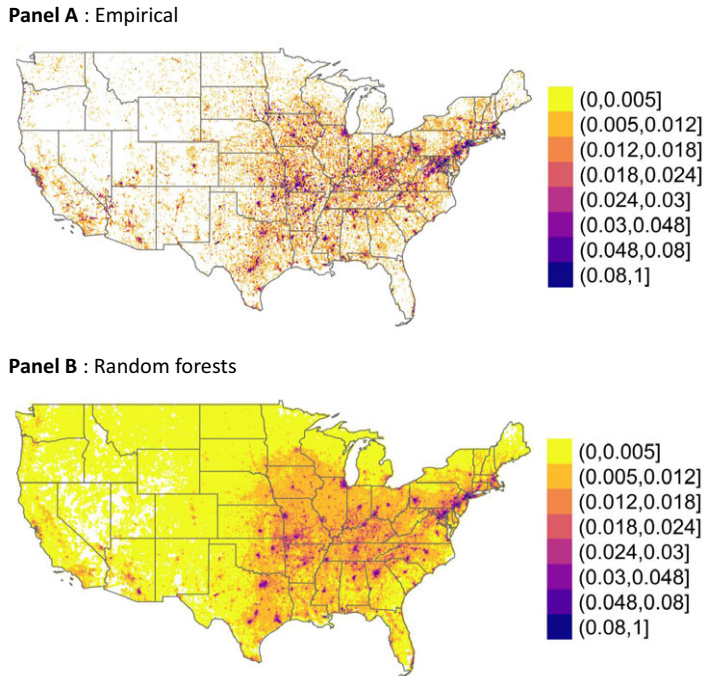
Panel A : Empirical

Panel B : Random forests

Figure 2. *Flood probabilities over the US: empirical (Panel A, top) versus predicted (Panel B, bottom) using the RF model (undersampling with 90% of zeroes, smallest set of covariates, and logged population).*
Empirical flood probability is calculated as the number of months with flood occurrence over the total number of months. A white cell means no occurrence has been observed. Predicted flood probabilities are computed as an average over months and years between 2007 and 2020.

was compiled. Such assessment is feasible since the predictors used (Section 3.1) are available globally or over North America, with the exception of population which covers the US only. In this case, we used the Gridded Population of the World (GPW) v4.11, which is also available from the SEDAC (CIESIN 2018).

We have also recalibrated the 15 models over the US, but only over regions whose Canadian KG climates are available in the US (therefore excluding grid cells whose climate would not contribute significantly to predicting flood dynamics in Canada such as Southern US states). As such, we show ROC and PR metrics with the claims data available from 2012 to 2020. The qualitative assessment compares time series of predicted flood probabilities between 2007 and 2020 knowing that major historical flood events took place in Toronto (2013, 2018) and Calgary (2013, 2019).

Table 2 shows the area under the ROC and PR curves for the 15 models applied in Canada. In bold face, we highlight the method with the highest metric for each line. We now see a different picture, as is usually the case when performing out-of-sample prediction exercises. The areas under the ROC curves now range within 0.77–0.90, which is lower than what we obtained over the US. That said, the performance is very good, particularly for the models with the smallest set of covariates and logged population, with metrics in the range of 0.83–0.90. It is particularly surprising to observe a value of 0.9 with the fifth model under the GLM; it appears that simpler specifications are performing well out-of-sample in Canada and that logged population captures the reality that urbanization beyond some level should not have the same impact on pluvial flood probability. We therefore have a solid case for the fifth set of models (smaller set of covariates, undersampling with 90% zeroes, with logged population density) which has the highest scores while being the fastest to fit (because of fewer covariates and

Table 2. *Area under the ROC and PR curves with flood claims from a Canadian insurer (2012-2020) for all 15 models considered. Note that "u/s" stands for undersampling.*

Models	ROC curve			PR curve		
	GLM	GAM	RF	GLM	GAM	RF
All, no u/s	**0.8136**	0.7812	0.7994	0.0153	0.0258	**0.0301**
All, 90% u/s	0.8065	0.7816	**0.8158**	0.0102	0.0236	**0.0339**
All, 50% u/s	0.7867	0.7811	**0.8147**	0.0077	0.0226	**0.0267**
Smaller, 90% u/s	0.7740	0.7874	**0.8276**	0.0094	0.0243	**0.0274**
Smaller, 90% u/s, log pop	**0.9009**	0.8219	0.8281	0.0152	0.0164	**0.0269**
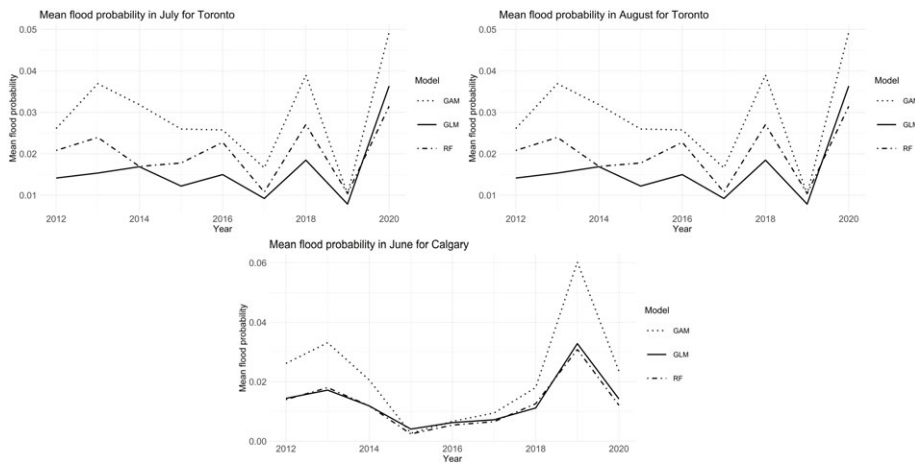
Figure 3. *Validation of pluvial flood models with predicted flood probabilities in Toronto over July and August (top row), and Calgary over June (bottom row) between 2012 and 2020. Models with the smallest set of covariates, 90% of zeroes, and logged population density were used.*

smaller sample size due to undersampling). As for the areas under the PR curves, RF have a slight advantage over the GAM.

We continue this section with a qualitative assessment of the models using selected flood events in Canada. Figure 3 shows major peaks in probabilities in July and August 2013 as well as in August 2018 which coincide with major flooding events in downtown Toronto. The 2013 floods in Toronto were among the most expensive for the insurance industry in Canada.

Looking at the bottom row for Calgary, we can distinguish significant peaks over June 2013 and June 2019. Although heavy rain triggered flooding in Calgary in June 2013, heavy snow accumulation upstream in the prior months magnified the intensity of the event, which might explain why the maximum observed in 2019 is higher than in 2013, because of the longer-term snow melt processes. Calgary also saw floods in June 2019 due to heavy rain from thunderstorms.

Moreover, in the three subplots of Figure 3, we observe that the GAM typically generates the largest range of flood probabilities, indicating that the model is the most responsive to changes in precipitation patterns. Finally, on the basis of the quantitative and qualitative validations, we find that the pluvial flood model fitted in the US provides strong predictions in Canada.

We conclude this section by providing a map of pluvial flood probabilities over Canada. Figure 4 shows that pluvial flooding is also concentrated in urban areas as was the case for the US. This is especially true in the Greater Vancouver area, Southern Quebec, and Ontario (including, Montreal and
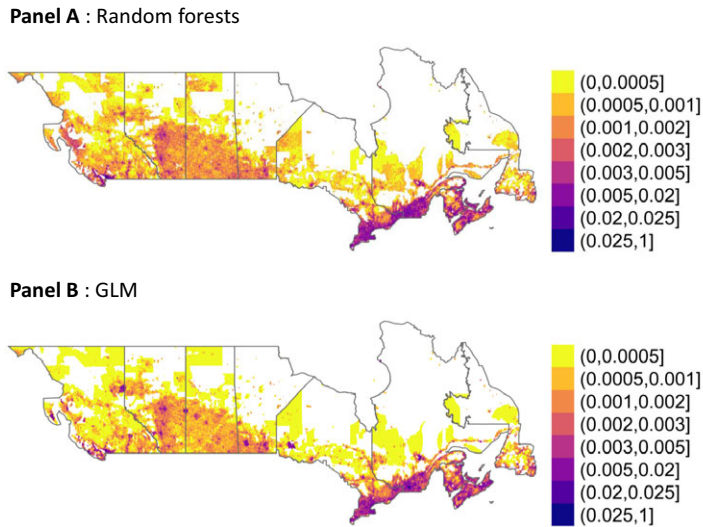
Panel A : Random forests

Panel B : GLM

Figure 4. *Predicted flood probabilities over Canada for the RF model (Panel A, top) and GLM (Panel B, bottom) using undersampling with 90% of zeroes, the smallest set of covariates, and logged population. Note that we cannot show historical flood probabilities to protect the confidentiality of the data. Similar plots for GAM are available in the SM.*

Toronto), as well as many urban areas of New Brunswick and Nova Scotia. Note that we cannot show historical claims patterns in Canada to protect the confidentiality of the clients. Moreover, we have more blank cells in Canada because more cells have no population (or too few) and because Northern Canadian climates do not exist in the US climate groups.

4. Future projections

In this section, we analyze pluvial flood probabilities predicted for the future. We first discuss the datasets used to build covariates (Section 4.1) and then how the statistical and machine learning methods have been applied with such covariates (Section 4.2). We conclude this section by analyzing the impacts of climate change and urbanization on pluvial flooding over the US, Canada, and for selected cities of both countries (Sections 4.3 and 4.4).

4.1 Data

4.1.1 Climate models outputs

We used climate model simulations from the Canadian Regional Climate Model version 5 (CRCM5) (Šeparović *et al.,* 2013; Martynov *et al.,* 2013), which is available from the Coordinated Regional Climate Downscaling experiment (CORDEX) – North America (NA) ensemble (World Climate Research Programme, WCRP). The six CRCM5 runs used are CCCma-CanESM2, MPI-ESM-LR, MPI-ESM-MR, UQAM-GEMatm-Can-ESMsea, UQAM-GEMatm-MPI-ESMsea, and UQAM-GEMatm-MPILRsea. The domain covers Canada and the US at a spatial resolution of 0.22° (about 25 km) from 1850 to 2100. The CRCM5 is known to well simulate realistic precipitation extremes, which is an important feature to model pluvial flooding (Martynov *et al.,* 2013; Martel *et al.,* 2020). Moreover, local dynamics obtained with RCMs are better than those obtained when applying statistical downscaling to GCMs (Maraun and Widmann, 2018).

We extracted data from 2007 to 2060 to match the initial date of the NOAA SED with an approximate 40-year future time horizon. Projections beyond 2060 are highly uncertain and heavily depend on the climate policies enacted today. All runs were forced with the RCP 8.5 scenario, which uses historical emissions before 2006 and scenario emissions in the future until 2100. Our analysis is focused on short-term projections (2010–2030, centered on 2020) and medium-term projections (2040–2060, centered on 2050). Although RCP 8.5 represents a pessimistic scenario, up to 2060 the concentration scenarios do not differ significantly.

Daily precipitation in the CRCM is expressed in $kg\,s^{-1}m^{-2}$ and temperature in degrees Kelvin. We multiplied precipitation by 86,400 to convert precipitation into mm/24 h and subtracted 273.15 to convert temperature into degrees Celsius.

4.1.2 Socioeconomic projections

Projections of future population are also available from the SEDAC (Jones and O'Neill, 2020). To test the sensitivity of results to changes in future population density, we used two population projections that are derived from the Shared Socioeconomic Pathways (SSP) scenarios from the IPCC AR6 (IPCC, 2022). We applied the SSP2 and SSP5 scenarios, which are respectively labeled as "Middle of the Road" and "Fossil-fueled Development" (O'Neill *et al.,* 2014; Fricko *et al.,* 2017; IPCC, 2022) which differ on population, GDP, and urbanization growth patterns. Both of these population projections are available on a $0.125° \times 0.125°$ grid and were reprojected to match the grids of the climate models. Note that in the IPCC Sixth Assessment Report, SSP2 and SSP5 are respectively tied to radiative forcings of 4.5 and 8.5 Wm^{-2}. Although combining the SSP2 and RCP 8.5 scenarios is useful for sensitivity analyses, the SSP2 socioeconomic projection is unable to generate the emissions and resulting radiative forcing of the RCP 8.5 scenario.

4.2 Methods

Because hourly precipitation is not available in the CRCM5 runs we analyzed, not all 15 models from Section 3.3 will be used for CCRA. Due to the unavailability of some variables, the predictive capability of the models, and computation times, we used the smaller set of covariates along with a target of 90% zeroes for each of the GLM, GAM, and RF. Covariates thus comprise 24-h precipitation, temperature, climate classification, five proportions of land use, and population density.

The first step consists in updating the covariates with output from the climate model. For each of the six runs of the CRCM5 over 2006 until 2060, we extract 24-h precipitation and compute the average daily temperature. Afterward, we record the maximum daily precipitation over the month and compute the average monthly temperature. The static variables such as Köppen–Geiger climate classifications and proportions of land use were held fixed until 2060 because no future projections were available. As for population density, we used two different assumptions: (1) SSP2 population projections until 2060 or (2) fixed from 2020 levels (SSP5 was also considered but results were not materially different over the time horizon considered). The latter thus fixes land use and urban extent and allows us to focus strictly on changes in future atmospheric conditions, whereas the former allows for interactions between increased urbanization and possibly more heavy rain.

The second step consists in computing flood probabilities with the updated covariates. For each month (12), year (54), and run (6) of the CRCM5, we computed flood probabilities using outputs of the CRCM5 as simulated covariates. We call these *simulated* flood probabilities, and they are available over the present and future climates. We interpret climate simulations over the present climate as alternate and plausible trajectories of the climate.

To mitigate the need to apply some kind of postprocessing (see Section 2.2) on flood probabilities, our analyses focuses on *differences* between two time periods (rather than looking at raw probabilities); therefore, assuming that any bias found in the CRCM5 over the historical period is likely to be of a

similar order in future projections. The SM provides an analysis of the CRCM5 over 2007–2020, and we find that such bias is very small in most areas.

4.3 Maps

Our analysis first compares flood probabilities between two time periods: 2010–2030 (centered on 2020) and 2040–2060 (centered on 2050), which are 30 years apart. We therefore average simulated probabilities across months, years, and runs of each time period for both the US (Figure 5) and Canada (Figure 6). Both figures therefore highlight the combined impacts of climate change and future urbanization on pluvial flooding hazard.

Whereas all three models agree that the West Coast of the US and Canada will be the most affected by changes in flood probabilities, there are however large discrepancies between predictions of the GAM/GLM and the RFs. The GAM/GLM families of models yield increases of pluvial flooding elsewhere in the US and Canada, concentrated in urban areas of Eastern US, Southern Quebec, and Ontario, whereas the RF shows close to no changes elsewhere. In fact, RF shows the smallest increases over the West Coast.

One should caveat the results from the RF models since RF and other tree-based methods are unable to extrapolate beyond the range of the training set (Hengl *et al.,* 2018). This becomes a major issue for CCRA because atmospheric variables such as precipitation and temperature patterns in the future may very well be different from their historical values.

4.4 Selected cities

We plot in this section the time series of the average (taken over months, grid cells of the city, and the six runs of the CRCM) annual simulated pluvial flooding probability for selected cities in the US (Figure 7) and in Canada (Figure 8). We observe increasing trends with different slopes across the different cities. Even though we averaged results over the six climate model runs, we still find substantial interannual variability. We also included in the SM similar plots for the GLM and RF. We find notably that there is nearly no trend with the RF and weak interannual variability, thus confirming the inability of the RF to extrapolate beyond the original training set.

In both Figures 7 and 8, we also isolated the effects of climate change from increased urbanization with the continuous and dotted lines. That is, the dotted line represents a scenario where population remains fixed after 2020, whereas the continuous line represents a scenario where population increases according to the SSP2 scenario. In the latter scenario, the population of New York, Chicago and Denver will continuously increase in the future, whereas Houston should see a population decrease from 2020 to 2030 and an increase thereafter. Although the increasing trend in flood probability seems primarily driven by changing patterns in temperature and precipitation, urbanization also plays an important role on flood hazard.

5. Portfolio applications

This section presents portfolio applications using the pluvial flood occurrence model under various hazard and exposure scenarios. This section has three objectives: (1) to demonstrate how the overall methodology could be used for CCRA; (2) to differentiate the impacts of changes in hazard and exposure and their interaction on portfolio losses, and (3) to illustrate the spatial heterogeneity of future climate and population projections.
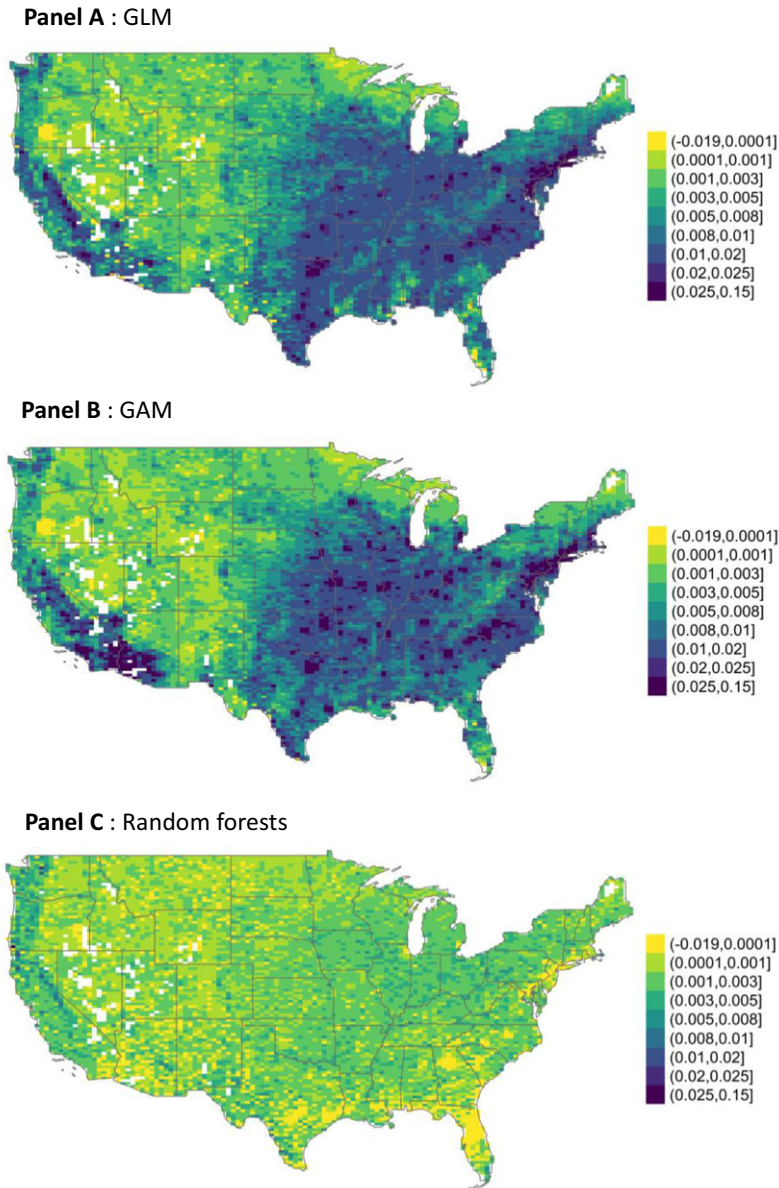
Panel A : GLM

Panel B : GAM

Panel C : Random forests

Figure 5. *Difference in simulated pluvial flood probability between 2040–2060 and 2010–2030 computed with the GLM (Panel A, top), GAM (Panel B, middle), and RF (Panel C, bottom) models over the US. Blank cells represent either too small population (in the past observations or future projections) or missing data.*

5.1 Methodology

The occurrence model applied to the CRCM5 and SSP2 population projection yields simulated flood probabilities for each grid cell, month and year between 2006 and 2060. We can therefore directly use these probabilities to simulate monthly flood occurrences over the future. Yet an important piece remains, linking flood occurrence to impact in terms of losses.
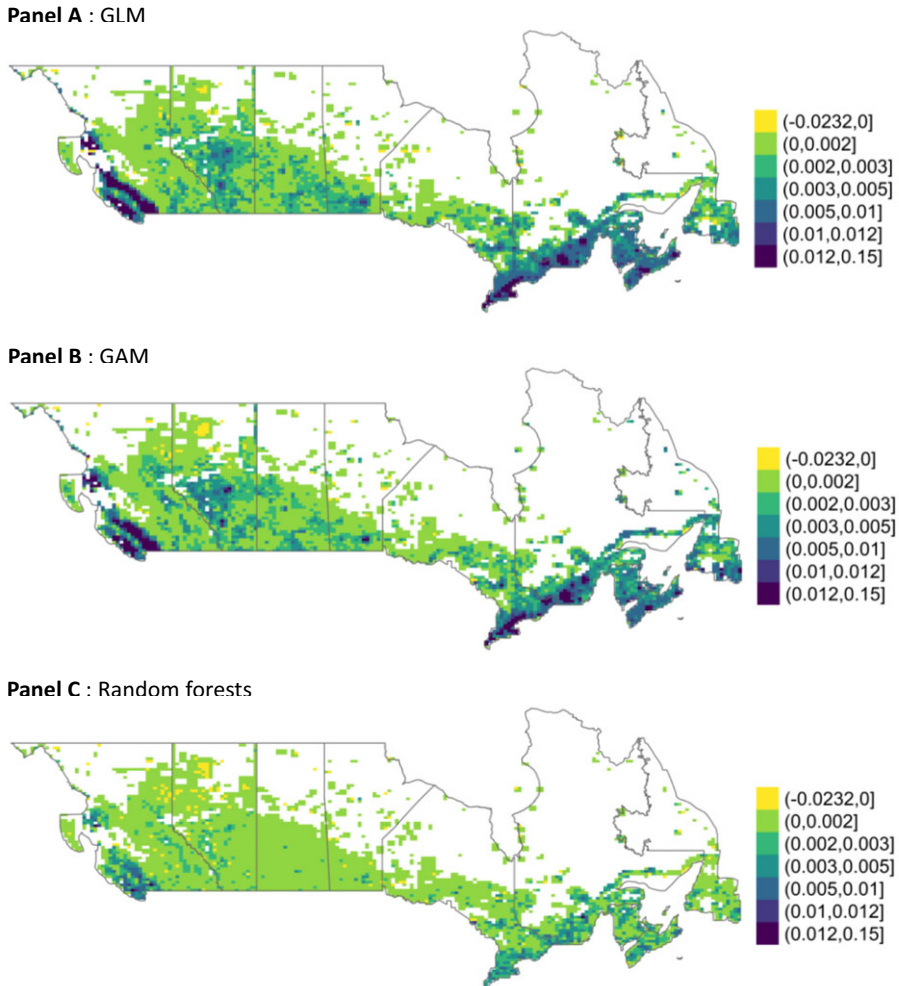
Panel A : GLM

Panel B : GAM

Panel C : Random forests

Figure 6. *Difference in simulated pluvial flood probability between 2040–2060 and 2010–2030 computed with the GLM (Panel A, top), GAM (Panel B, middle), and RF (Panel C, bottom) models over Canada. Blank cells represent either too small population (in the past observations or future projections) or missing data.*

The typical flood risk modeling chain (see e.g., Boudreault *et al.,* 2020 for fluvial flooding or Figure 1 for a general setup) entails using the location (latitude and longitude) and characteristics of each building in combination with a damage model (e.g., a damage curve linking water depth and dollars of losses) to represent losses from flooding. But with hazard information available on a 10–25 km grid in the form of flood occurrence instead of water depth, we do not aim to analyze impacts at the street level, and as such, the exact location of each building is not necessary for this exercise. For similar reasons, we will also ignore the vulnerability of each building and rather assume that each flooded property suffers a fixed or random loss amount. Aggregating exposure value, or the number of households insured per grid cell at a resolution similar than that of the climate model is straightforward for an insurer. But for this paper, we will rather build generic insurance portfolios based on the population data described in Sections 3.1 and 4.1.

It remains to determine the number of homeowners that are flooded when there is flood occurrence in a given grid cell. One can fix that number as a given percentage, but we instead modeled it as a
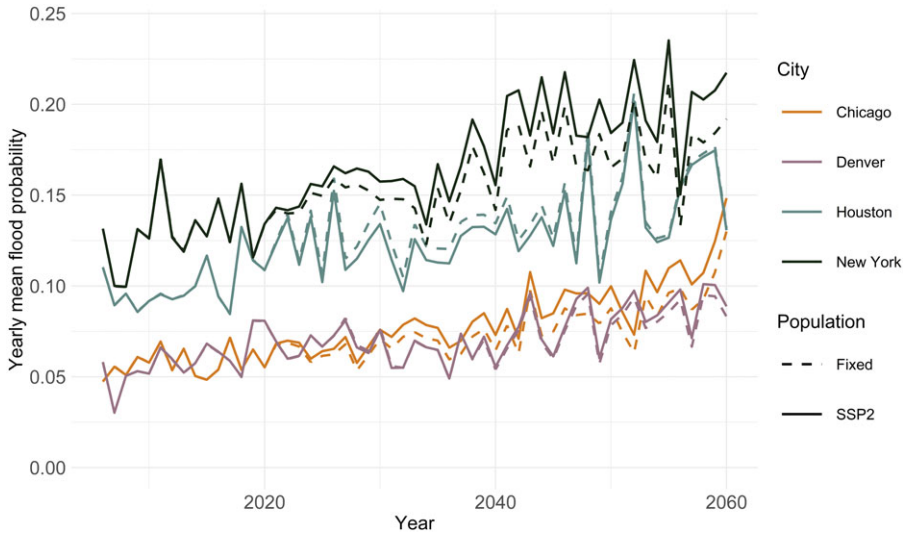
Figure 7. *Annual simulated pluvial flood probability from 2006 to 2060 over New York, Houston, Chicago and Denver with the GAM model. Similar plots for GLM and RF are available in the SM.*
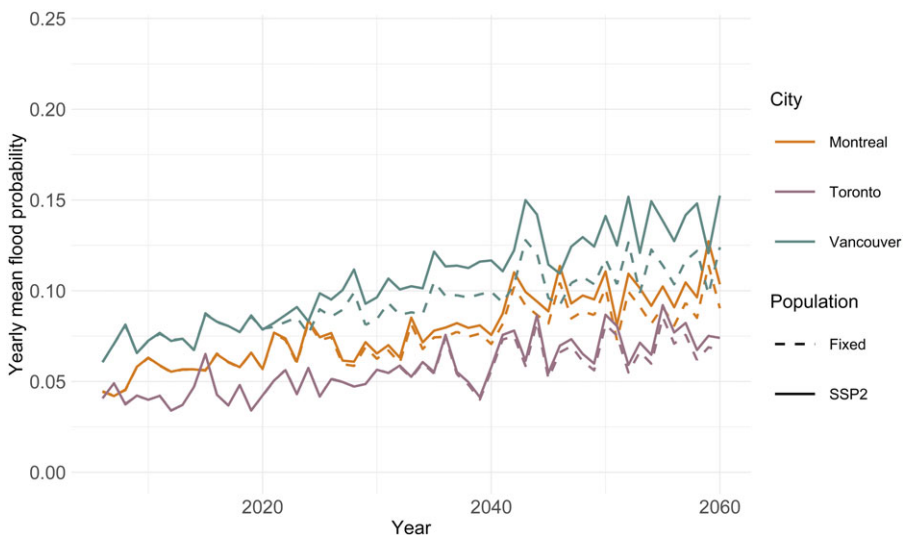

Figure 8. *Annual simulated pluvial flooding probability from 2006 to 2060 over Montreal, Toronto and Vancouver with the GAM model. Similar plots for GLM and RF are available in the SM.*

beta-distributed random variable with fixed mean and a fixed upper percentile. As such, this adds randomness as to how extreme precipitation may locally affect a community, replicating the effects of spatial contagion within each grid cell.

The specific methodology is as follows. We have split the time horizon until 2060 into two time periods: present climate (2010–2030), centered around 2020, and future climate, centered around 2050 (2040–2060). A 30-year time horizon is reasonable for an insurer for strategic decision-making and solvency analyses, while avoiding the considerable uncertainty tied to considering climate up to 2100, which is heavily dependent on current climate policies. Each model run (6) and year (20) within each

time period is assumed to be independent and identically distributed. This yields 120 climate simulations under the present climate and another 120 climate simulations under the future climate.

For the present climate, we draw 10,000 random numbers in order to randomly select a climate from the 120 available. For the selected climate, we compute simulated flood probabilities for each grid cell and month over Canada and the US. We then draw Bernoulli random variates according to these probabilities over each grid cell assuming that flood occurrences conditional upon the climate is spatially independent. If there is a flood in a given grid cell, we then randomly draw from a beta distribution with mean 2% and 99-th percentile equal to 20% to represent the percentage of homeowners affected. For each household affected by a flood, we assume a loss of $25,000. The value of $25,000 is based on the average damage given a pluvial flood per property, whereas 2% is selected to replicate industry losses. These choices do not make a material difference to understanding the *relative* impacts of climate change.

The previous steps were then repeated for the future climate as well. In both cases, we worked with the GAM and the smaller set of covariates fitted with a targeted 90% of zeroes. GLM yields similar results, whereas the RF has been excluded for the reasons explained in Section 4.

5.2 Results

To meet the objectives presented at the beginning of this section, we construct three scenarios for future change in the hazard and exposure. The baseline scenario represents our best estimate of the current loss distribution. It is based on present-day hazard (2020) and exposure (2020). The second scenario assesses the sensitivity of the insurer's current exposure (2020) to changes in hazard (2050), including future projections of temperature, precipitation, and urbanization. It represents what would be typically asked for reporting and regulating purposes to assess the impacts of future pluvial flood hazard. In this case, the insurer's portfolio is held fixed in the future, as if the insurer would not underwrite additional risks. Finally, the third scenario includes changes in both the hazard (2050) and exposure (2050), depicting a situation where a company underwrites in a similar manner and fixes its future market share instead. Such a scenario also highlights possible interactions between hazard and exposure where population could increase or decrease in riskier or safer areas. In all cases, we fixed the market share to 100% of the corresponding geographic region, which therefore proxies industry losses. Note that it is possible to create a fourth scenario in which hazard is fixed as of 2020 but exposure is that of 2050. Of lesser interest for insurance applications, this fourth scenario would support understanding the contribution of projected exposure to future flood risk. This is left for future research.

Table 3 shows the results of the latter three scenarios for four portfolios, fully underwritten in Quebec, Ontario, Canada, or over the US. Note that $ amounts have not been adjusted for inflation and reflect losses as of 2020. All risk measures were computed with 10,000 simulations, whereas the mean and standard deviations were validated with closed-form expressions that are straightforward to derive. The SM shows the equivalent of Table 3 but for each of the 10 Canadian provinces and the 10 most populous US states.

With the four portfolios illustrated in Table 3, we see that even in the aggregate, changes in hazard can be profoundly different across regions. Under the second scenario (hazard of 2050 but exposure of 2020), losses are expected to increase by nearly 50% in both Quebec and Ontario, whereas the increase is lower Canada-wide (about +40%) or in the US (+30%). Expressed differently, such increases represent 0.88% to 1.36% per year when compounded annually. Across states and provinces, SM Section 3 shows more homogeneity across Canadian provinces (increases of about 40–50% with the exception of BC and PEI) than in the US, where increases range from 15 to 50%.

Although we find benefits to diversification, country-wide effects of climate change on pluvial flood are expected to be relatively more significant in Canada than in the US. According to ECCC (Bush and Lemmen, 2019), the average temperature increase in Canada is expected to be greater than in the US. This indicates that the air over Canada, being warmer, could hold more moisture that would in turn drive more intense rainfall, all other things being equal.

Table 3. *Portfolio loss statistics for four portfolios and three scenarios for changes in hazard and exposure (in millions of 2020 dollars). Relative difference in % shown between parentheses (compared to the baseline scenario).*

	Hazard	Exposure	Average	Std. dev.	90th perc.	95th perc.	99th perc.
Quebec	2020	2020	471	653	1154	1707	3291
	2050	2020	692 (47%)	767 (18%)	1601 (39%)	2234 (31%)	3872 (18%)
	2050	2050	913 (94%)	995 (53%)	2067 (79%)	2947 (73%)	5070 (54%)
Ontario	2020	2020	693	827	1613	2291	4281
	2050	2020	1029 (49%)	987 (19%)	2243 (39%)	3059 (34%)	4827 (13%)
	2050	2050	1285 (85%)	1175 (42%)	2713 (68%)	3701 (62%)	5842 (36%)
Canada	2020	2020	1834	1260	3518	4327	6230
	2050	2020	2605 (42%)	1462 (16%)	4562 (30%)	5467 (26%)	7583 (22%)
	2050	2050	3240 (77%)	1717 (36%)	5469 (55%)	6600 (53%)	9219 (48%)
USA	2020	2020	18,840	5666	26,660	28,948	33,057
	2050	2020	24,499 (30%)	6625 (17%)	32,551 (22%)	35,165 (21%)	42,551 (29%)
	2050	2050	29,775 (58%)	8261 (46%)	40,112 (50%)	43,508 (50%)	52,036 (57%)

In the third scenario, both the hazard and exposure change in the future. There are however nontrivial interactions between changes in hazard and exposure depending on where population will live. Indeed, if for example current and future population move to areas with increasing hazard, then portfolio losses will increase at a faster pace than population growth. Table 3 shows that the third scenario yields losses much greater than the second scenario with significant heterogeneity. For example, the Quebec portfolio losses nearly double, while the US-wide portfolio losses increase by about 60%. Expressed on an annual basis, the compounding effects of increasing hazard and exposure mean that losses should increase by a rate of 1.6–2.2% annually. Across states and provinces, SM Section 3 shows variations between 40 and 95% over Canadian provinces and the top 10 US states, which is significant.

Significant trends in future losses should not be a surprise. Adding inflation of about 3% (which is slightly above the historical inflation over the last 40 years, but still below the inflation observed in 2022–2023) could yield a compound annual rate of increase in losses of over 5% (all else being equal, namely adaptation). Note that according to the Parliamentary Budget Officer in Canada, claims to the Disaster Financial Assistance Arrangements due to flooding have quadrupled over the last 40 years (Office of the Parliamentary Budget Officer, 2016), and therefore, the figures we present are clearly not unrealistic.

We conclude this section by analyzing the (kernel-smoothed) loss distributions in each of the three scenarios for the four portfolios. We clearly see rightward shifts in Figure 9 (as well as in the log-log version of Figure 9 provided in the SM) as we move from the baseline scenario (hazard of 2020, exposure of 2020) to the third scenario (hazard of 2050, exposure of 2050). The Quebec and Ontario portfolios are right-skewed and heavy-tailed, even more so than the Canadian and US portfolios. Judging by the upper percentiles, there does not appear to be a significant thickening of the right tail under the third scenario, but this is based on 10,000 simulations founded on 120 different climates, which could limit the potential to capture extreme losses. The conditional independence assumption may have played a role in limiting the potential for extreme portfolio losses; a longer discussion is included in Section 6. Therefore, we caveat interpreting and extrapolating future impacts of climate change on extreme portfolio losses based on Table 3 and Figure 9.

This portfolio application shows the value for (re)insurance companies to invest in better underwriting practices and/or work with communities to attenuate the financial impacts of climate change on flooding. It also highlights the importance of quantitative CCRA to support such strategic decision-making at the organization level.
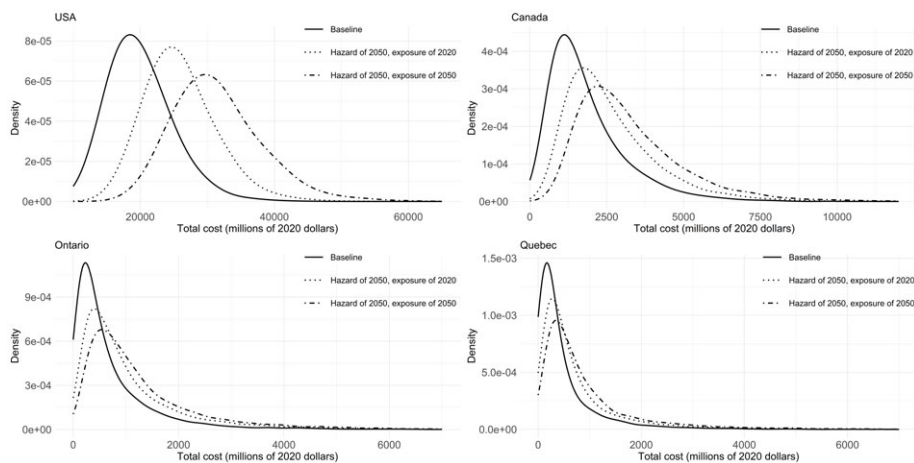
Figure 9. *Probability density functions of portfolio losses for each portfolio and scenario.*

6. Discussion and conclusion

As reporting and regulatory requirements evolve, actuaries will increasingly need to factor in climate change into various business functions such as underwriting, reserving, and strategic decision-making. Climate risks are not new to actuaries, but climate change might force the actuarial profession to not only look for answers in past data but also look forward in the future using climate models. Integrating climate models into actuarial assessments is new to the profession, and this paper has shown that CCRA can also be viewed as a data science problem. This is an important outcome given the talent pool that insurers typically recruit from.

One objective of the paper was to assess how pluvial flood risk may affect an insurance portfolio in the future. Using historical data on pluvial flood occurrences, we applied statistical and machine learning methods to better understand the relationship between these flood occurrences and atmospheric and socioeconomic variables (fitting and validation step). Using climate model outputs as simulations of atmospheric variables over the present and future, we then computed pluvial flood probabilities over Canada and the US until 2060 (projection and simulation step). Finally, with a simple portfolio model founded on a single flood occurrence model, we evaluated how changes in hazard and exposure may impact different insurance portfolios. The overall approach depicted in the paper is designed for large-scale applications that do not necessarily require street-level information, as is often the case for scenario and trend analyses. There is obviously a trade-off between speed, flexibility, cost, and precision for all applications, and the methodology described here is no exception.

We found that standard statistical and machine learning methods such as GLM, GAM, and RF are very good at predicting pluvial flood occurrence over the US and that such fits yield solid predictive skill out-of-sample over Canada. We used six runs of the CRCM5 available in the CORDEX-NA ensemble to compute flood probabilities over the US and Canada. We found strongly heterogeneous impacts of climate change over urban areas in Canada and the US. Results are consistent whether we use GLM or GAM to explain the link between atmospheric variables and flood occurrences, but RF are clearly not recommended for CCRA due to their inability to make reliable predictions outside of their training domain. Predicted flood probabilities from the RF for future climates go against the mounting evidence that climate change will increase heavy rain episodes and pluvial flooding (Bush and Lemmen, 2019; IPCC, 2022).

There are several areas worth investigating for future research. First, CCRA yields many uncertainties that stem from the natural variability of climate, the complexity of natural hazards, and also the unpredictable future climate policies and resulting GHG emissions. To assess the size and impact of such

uncertainties, one approach is to evaluate the sensitivity of the results to different emissions scenarios (RCPs and SSPs) and different classes of models (higher resolution GCMs from the CMIP6 ensemble used in the AR6 of the IPCC). With a sample made of 168 grids of 143,922 cells or 2.4 million observations, it would be interesting to evaluate the ability of artificial neural networks (see, e.g., Wüthrich, 2018; Richman and Wüthrich, 2021; Chen *et al.,* 2023, as well as long short-term memory (LSTM) that are popular in hydrology, see e.g., Kratzert *et al.,* 2018) and other deep learning methods to extrapolate out-of-sample over Canada and over future climates.

There are some limitations to the flood occurrence models that could be alleviated in future research. Namely, the conditional independence assumption is a somewhat strong assumption made to use traditional statistical and machine learning methods for classification problems. It results that grid cells are spatially dependent because of, for example, atmospheric variables, but there is no mechanism to create contagion other than from the covariates. In other words, if one grid cell floods, this does not modify the flood probability of contiguous grid cells. That said, the extent to which conditional independence affects spatial diversification and the potential for extreme aggregate losses in fully diversified portfolios over Canada or the US remains to be investigated. Extending the statistical models by allowing contagion over contiguous grid cells accounting for local and nearby topography would be an interesting approach to answer such question in the context of pluvial flooding. Finally, lack of data on the intensity of pluvial floods, expressed either in terms of return period, water depth or speed, prevented us from modeling the severity of pluvial flooding. This may in turn result in underestimated potential losses. Future research could look into jointly dependent frequency and intensity models for present and future pluvial flooding.

References

Bank of England. (2019) *The 2021 biennial exploratory scenario on the financial risks from climate change.*

Beck, H.E., Wood, E.F., Pan, M., Fisher, C.K., Miralles, D.G., van Dijk, A.I.J.M., McVicar, T.R. and Adler, R.F. (2019) MSWEP V2 global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *Bulletin of the American Meteorological Society*, **100**(3), 473–500. ISSN: 0003-0007, 1520-0477, accessed July 29, 2021. https://doi.org/10.1175/BAMS-D-17-0138.1. https://journals.ametsoc.org/view/journals/bams/100/3/bams-d-17-0138.1.xml.

Blier-Wong, C., Cossette, H., Lamontagne, L. and Marceau, E. (2020) Machine learning in P&C insurance: A review for pricing and reserving. *Risks*, **9**(1), 4.

Boudreault, M., Grenier, P., Pigeon, M., Potvin, J.-M. and Turcotte, R. (2020) Pricing flood insurance with a hierarchical physics-based model. *North American Actuarial Journal*, **24**(2), 251–274.

Bush, E. and Lemmen, D.S. (2019) *Canada's Changing Climate Report.* Technical report. Government of Canada.

Canada, Public Safety. (2022a) *Canadian Disaster Database.* https://www.publicsafety.gc.ca/cnt/rsrcs/cndn-dsstr-dtbs/index-en.aspx.

Canada, Public Safety. (2022b) *Floods*, August. https://www.publicsafety.gc.ca/cnt/mrgnc-mngmnt/ntrl-hzrds/fld-en.aspx.

Carozza, D.A. and Boudreault, M. (2021) A global flood risk modeling framework built with climate models and machine learning. *Journal of Advances in Modeling Earth Systems*, **13**(4), e2020MS002221.

CEC, Commission for Environmental Cooperation. (2015) *Land Cover 30m, 2015 (Landsat and RapidEye)* [in en-US]. Publication Title: Commission for Environmental Cooperation. Accessed August 3, 2021. http://www.cec.org/north-american-environmental-atlas/land-cover-30m-2015-landsat-and-rapideye/.

Chen, D., Rojas, M., Samset, B.H., Cobb, K., Diongue Niang, A., Edwards, P., Emori, S., et al. (2021) Framing, context, and methods. In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (eds. V. Masson-Delmotte, P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, et al.), pp. 147–286. Cambridge, UK and New York, NY, USA: Cambridge University Press. https://doi.org/10.1017/9781009157896.003.

Chen, Z., Lu, Y., Zhang, J. and Zhu, W. (2023) Managing weather risk with a neural network-based index insurance. *Management Science*.

CIESIN. (2017) U.S. Census Grids (summary file 1), 2010. *Center for International Earth Science Information Network, Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC)*.

CIESIN. (2018) Gridded Population of the World, version 4.11 (GPW v4.11): Population Count, revision 11. *Center for International Earth Science Information Network, Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC)*.

Cook, N.R. (2007) Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*, **115**(7), 928–935.

Davies, R., Behrend, J. and Hill, E. (2021) *FloodList*. https://floodlist.com/data-api.

FEMA. (2017) Flood Insurance Reform: FEMA's Perspective. Technical report. Federal Emergency Management Agency, March.

Feng, B., Zhang, Y. and Bourke, R. (2021) Urbanization impacts on flood risks based on urban growth data and coupled flood models. *Natural Hazards*, **106**(1), 613–627.

Financial Stability Board. (2017) *Final Report: Recommendations of the Task Force on Climate-related Financial Disclosures*.

Fricko, O., Havlik, P., Rogelj, J., Klimont, Z., Gusti, M., Johnson, N., Kolp, P., Strubegger, M., Valin, H., Amann, M., et al. (2017) The marker quantification of the Shared Socioeconomic Pathway 2: a middle-of-the-road scenario for the 21st century. *Global Environmental Change*, **42**, 251–267.

Ganganwar, V. (2012) An overview of classification algorithms for imbalanced datasets. *International Journal of Emerging Technology and Advanced Engineering*, **2**, 42–47.

Guha-Sapir, D., Below, R. and Hoyois, P. (2022) *EM-DAT: the CRED/OFDA International Disaster Database*. http://www.emdat.be.

Hengl, T., Nussbaum, M., Wright, M.N., Heuvelink, G.B.M. and Gräler, B. (2018) Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ*, **6**, e5518.

ICLR. (2021) *Focus on Types of Flooding*. Technical report. Institute for Catastrophic Loss Reduction, April.

IPCC. (2021a) Annex VII: Glossary [Matthews, J.B.R., V. Möller, R. van Diemen, J.S. Fuglestvedt, V. Masson-Delmotte, C. méndez, S. Semenov, A. Reisinger (eds.)] In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, (ed. V. Masson-Delmotte, P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, et al.), pp. 2215–2256. Cambridge, UK and New York, NY, USA: Cambridge University Press. https://doi.org/10.1017/9781009157896.022.

IPCC. (2021b) *Hazard, vulnerability, exposure and risk*. https://www.ipcc.ch/srocc/chapter/technical-summary-ts-0-introduction/ipcc-srocc-ts_4/.

IPCC. (2022) Summary for policymakers. In *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (ed. H.O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, et al.). In Press. Cambridge, UK: Cambridge University Press.

James, G., Witten, D., Hastie, T. and Tibshirani, R. (2021) *An Introduction to Statistical Learning with Applications in R*, 2nd edition. Springer.

Jin, Z. and Erhardt, R.J. (2020) Incorporating climate change projections into risk measures of index-based insurance. *North American Actuarial Journal*, **24**(4), 611–625.

Jones, B. and O'Neill, B.C. (2020) *Global One-eighth Degree Population Base Year and Projection Grids Based on the Shared Socioeconomic Pathways, Revision 01*. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).

Kratzert, F., Klotz, D., Brenner, C., Schulz, K. and Herrnegger, M. 2018. Rainfall–runoff modelling using long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences*, **22**(11), 6005–6022.

Maraun, D. and Widmann, M. (2018) *Statistical Downscaling and Bias Correction for Climate Research*. Cambridge University Press.

Martel, J.-L., Mailhot, A. and Brissette, F. (2020) Global and regional projected changes in 100-yr subdaily, daily, and multiday precipitation extremes estimated from three large ensembles of climate simulations. *Journal of Climate*, **33**(3), 1089–1103.

Martynov, A., Laprise, R., Sushama, L., Winger, K., Šeparović, L. and Dugas, B. (2013) Reanalysis-driven climate simulation over CORDEX North America domain using the Canadian Regional Climate Model, version 5: Model performance evaluation. *Climate Dynamics*, **41**, 2973–3005.

Mitchell-Wallace, K., Jones, M., Hillier, J. and Foote, M. (2017) *Natural Catastrophe Risk Management and Modelling: A Practitioner's Guide*. Oxford, UK: John Wiley & Sons.

Mosavi, A., Ozturk, P. and Chau, K.-w. (2018) Flood prediction using machine learning models: Literature review. *Water*, **10**(11), 1536.

Muschelli, J. III. (2020) ROC and AUC with a binary predictor: A potentially misleading metric. *Journal of Classification*, **37**(3), 696–708.

National Oceanic and Atmospheric Administration, Office of Oceanic and Atmospheric Research, Physical Sciences Laboratory, and Earth System Research Laboratories. (2021) *CPC Global Temperature data.* Accessed August 3, 2021. https://psl.noaa.gov/data/gridded/data.cpc.globaltemp.html.

NOAA, National Oceanic and Atmospheric Administration. (2021) *Storm Events Database \| National Centers for Environmental Information*. Accessed July 29, 2021. https://www.ncdc.noaa.gov/stormevents/.

O'Neill, B.C., Kriegler, E., Riahi, K., Ebi, K.L., Hallegatte, S., Carter, T.R., Mathur, R. and Van Vuuren, D.P. (2014) A new scenario framework for climate change research: The concept of shared socioeconomic pathways. *Climatic Change*, **122**, 387–400.

Office of the Parliamentary Budget Officer. (2016) *Estimate of the Average Annual Cost for Disaster Financial Assistance Arrangements due to Weather Events*. https://www.pbo-dpb.gc.ca/web/default/files/Documents/Reports/2016/DFAA/DFAA_EN.pdf.

OSFI. (2023) *Guideline B-15*. Technical report. Office of the Superintendent of Financial Institutions, March.

Peel, M.C., Finlayson, B.L. and McMahon, T.A. (2007) Updated world map of the Köppen-Geiger climate classification [in English]. *Hydrology and Earth System Sciences*, **11**(5), 1633–1644. ISSN: 1027-5606, accessed August 3, 2021. https://doi.org/10.5194/hess-11-1633-2007. https://hess.copernicus.org/articles/11/1633/2007/hess-11-1633-2007.html.

Richman, R. (2021a) AI in actuarial science–a review of recent advances–part 1. *Annals of Actuarial Science*, **15**(2), 207–229.

Richman, R. (2021b) AI in actuarial science–a review of recent advances–part 2. *Annals of Actuarial Science*, **15**(2), 230–258.

Richman, R. and Wüthrich, M.V. (2021) A neural network extension of the Lee–Carter model to multiple populations. *Annals of Actuarial Science*, **15**(2), 346–366.

Ross, B. and Bruce, B. (2010) *Canadian Meteorological Centre (CMC) Daily Snow Depth Analysis Data, Version 1.* Accessed August 3, 2021. https://doi.org/10.5067/W9FOYWH0EQZ3. http://nsidc.org/data/NSIDC-0447/versions/1.

Saerens, M., Latinne, P. and Decaestecker, C. (2002) Adjusting the outputs of a classifier to new a priori probabilities: A simple procedure. *Neural Computation*, **14**(1), 21–41. ISSN: 08997667, accessed December 9, 2021. https://doi.org/10.1162/089976602753284446.

Saito, T. and Rehmsmeier, M. (2015) The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PloS One*, **10**(3), e0118432.

Seirup, L. and Yetman, G. (2006) U.S. Census Grids (Summary File 1), 2000. *Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC)*.

Šeparović, L., Alexandru, A., Laprise, R., Martynov, A., Sushama, L., Winger, K., Tete, K. and Valin, M. (2013) Present climate and climate change over North America as simulated by the fifth-generation Canadian regional climate model. *Climate Dynamics*, **41**(11-12), 3167–3201.

UN, (United Nations). (2023) *Disaster Risk Management, UN-SPIDER Knowledge Portal.* https://www.un-spider.org/risks-and-disasters/disaster-risk-management.

UNDRR. (2017) *Report of the open-ended intergovernmental expert working group on indicators and terminology relating to disaster risk reduction.* Technical report. United Nations Office for Disaster Risk Reduction, February.

Wüthrich, M.V. (2018) Neural networks applied to chain–ladder reserving. *European Actuarial Journal*, **8**, 407–436.

Wüthrich, M.V. (2023) Model selection with gini indices under auto-calibration. *European Actuarial Journal*, **13**(1), 469–477.

Yeo, N., Lai, R., Ooi, M.J. and Liew, J.Y. (2019) *Literature Review: Artificial Intelligence and Its Use in Actuarial Work*. Technical report. Society of Actuaries.