

2. Materials and Methods

2.1. Data Sources and Scales

The hazard data we use to develop our adaptation model refer spatially to 395,526 U.S. Census blocks that intersect the CRB (Figure 1). These blocks belong to seven different U.S. states and 89 counties, as presented in Figure 2. In comparison to larger administrative spatial units (i.e., counties), which are usually the ones referred to in the literature, Census blocks provide an increased level of precision for the identification and prediction of extremes and for the attribution of damage. Summary statistics for the different possible spatial scales of administrative units intersecting CRB are presented in Table 1. Census blocks considered have an average size of about 2 km² and a median area of only 0.04 km².

Figure 1. U.S. Census blocks in the Colorado river basin. The 395,526 U.S. Census blocks (represented in red) intersecting CRB (represented in black—U.S. Census 2010 boundaries [38]).

Figure 2. The Colorado river basin and main human scales. U.S. Census counties (represented in black) and states (from left to right, from bottom to top: California in light blue, Arizona in olive, New Mexico in blue, Nevada in pink, Utah in dark green, Colorado in orange, and Wyoming in green) intersecting CRB (represented in black—U.S. Census 2010 boundaries [38]).

Table 1. Summary statistics for main U.S. Census units intersecting the CRB (areas in km²).

Our economic model is calibrated using data for the 1997–2017 (current) period. The hazard data we use for analyzing paste hazard expectations extend from 1977 to 2017. Future scenarios use hazard and damage data predicted for the period 2018–2100. The data we use are scaled to quarterly time intervals.

2.1.1. Hazard

A natural hazard value is a quantification of the intensity of the extreme considered. Here, we define hazards using the concurrent estimation of multiple indicators of extreme events for flooding. All data are derived from simulations using the Variable Infiltration Capacity (VIC) hydrologic model [39], which is implemented as described in Bennett et al. [30]. VIC forcing data (i.e., temperature, precipitation, and wind speed) for historical simulations are generated from historical gridded climate data [40], while future climate data are generated from downscaled projections from the Multivariate Adaptive Constructed Analogue (MACA) database [41]. VIC parameters (i.e., soils, vegetation, and topography) are described in detail in Bennett et al. [30]. VIC is calibrated using a multiobjective, automated calibration tool to naturalized streamflow for select locations throughout CRB [30,42].

For this work, we used synoptic time scale (maximum daily values extracted for 5 days periods, with 73 5-day intervals in each year) data for each of three extreme flooding indicators following the work described in Bennett et al. [29]: maximum soil moisture in the entire soil column, maximum precipitation, and maximum runoff. We consider a historical period (1970–1999) along with current (2000–2017) and multiple future scenarios for different times (2020–2039,

2040–2069, and 2070–2099).

Historical and future scenarios are based on results from three different Earth System Models (ESMs): IPSL-CM5A-LR, GFDL-ESM2G, and MIROC-ESM. These three ESMs are selected as they represented a range of moderate 2040–2069 conditions (calculated as a difference from the 1970–1999 period) for CRB across the ESMs available in the MACA data set. Figure 3 illustrates 2040–2069 and 2070–2099 periods for all ESMs available in MACA. ESMs selected for this work include a warm/dry scenario (IPSL-CM5A-LR) and two moderately wet scenarios (warm/wet, GFDL-ESM2G and hot/wet, MIROC-ESM, Table S1). It is worth noting that IPSL-CM5A-LR tends towards drier and hotter, while the other two ESMs remain in the moderate warm/hot and wet range of the ESMs towards the end of the century, which has implications for the flooding results presented herein.

Figure 3. Annual average range of temperature in °C and precipitation in percentage values for two time periods for the three ESM selected for study calculated based on the difference from the historical period (1970–1999).

Synoptic data are spatially averaged for each Natural Resource Counsel (NRCS), United States Department of Agriculture (USDA) hydrological cataloguing unit, level eight (HUC8) watersheds for the CRB (134 total). Figure 4 illustrates spatially averaged synoptic data for each flooding indicator for a single HUC8 watershed: the Uncompahgre River basin. Concurrent copulas are calculated using the three flooding indicators following methods described below for each of the HUC8 basins, based on historical VIC simulation.

Figure 4. Input values to the copulas for the historical synoptic time series of maximum precipitation (precx, mm), maximum soil moisture (soilmx, mm), and maximum streamflow (qx, mm) for IPSL-CM5A-LR ESM for the Uncompahgre River basin in the Upper Colorado. The 95th percentile of the values is illustrated with dashed lines. Histograms show the distribution of the raw data values with lines for marginal fits.

The multidimensional representation of the joint distributions of relevant hydrological climate impacts is based on the concept of statistical copulas [43]. If X is a random vector of m random variables X_1, \dots, X_m

(e.g., streamflow, temperature, and evapotranspiration), the copula (CX) gives their joint distribution according to the following

(Sklar's theorem):

$$FX(x) = CX(FX_1(x_1), \dots, FX_m(x_m))$$

(1)

where the $FX_1(x_1), \dots, FX_m(x_m)$ functions are the marginal cumulative distribution functions of the individual variables, and $CX: [0, 1] \times \dots \times [0, 1] \rightarrow$

$[0, 1]$.

$[0, 1]$.

The estimation of the concurrent distribution consists of finding a suitable multidimensional representation of (possibly) correlated data. The process of finding such representation takes place in two stages. The marginal univariate distributions for each individual variable is found, and the best model is chosen from goodness of fit (gof) tests by scaling and transforming raw data values as needed (see Figure 4). We use the R package `fitdistrplus`, which supports a

number of probability distributions such as normal, gamma, normal, log-normal, exponential, gumbel, and beta. The 'fitdist' function is applied to fit the data, and 'gofstat' is used to perform gof tests. Then, we obtain the joint distribution by fitting a copula to the selected marginal distributions.

Copulas are selected from two popular copula families that included Elliptical copulas, which are restricted to radial symmetries and a more flexible family called Archimedean copulas. Typical Elliptical copulas include Gaussian and t-student copulas. Among the Archimedean copulas, the most common are Clayton, Frank, Gumbel, and Joe. To select the final copulas to represent flooding across the three dimensions, we iterate through a series of copulas to find the best fit using a quasi-Newton method [44].

Concurrent extremes values are transformed both spatially and temporally to match the scale of the adaptation model data. We first disaggregate copula results within the 95th percentile of the copula's range and then identify when these ranges were met or exceeded in the current and future data at the grid cell scale. This results in a conservative estimate of the changes in extremes, which we felt was most prudent provided the range of uncertainties present within the range of ESM estimates [45]. Then, we use spatial boundaries defining the U.S. Census blocks to query grid-cell scale copula results and select the spatial cells intersecting each block. As a last step, we compute the average value of flooding in those cells for each quarter considered. The final data set contains an average value of floodings for each block, quarter, and ESM considered. Copula results for the historical and future for the Uncompaghre River basin are shown as an example in Figure 5.

Figure 5. Historical (a) and future (b) copula for high precipitation (>1 of standard deviation) plotted as soil moisture maximum against maximum runoff for the Uncompaghre River basin. Darker colors indicate a higher density of data. Dashed lines show 95th, 90th, and 75th percentiles of the data.

2.1.2. Exposure

Exposure refers to the monetary value of human assets and activities present in a specific space and time. Several different approaches can be adopted for quantifying exposure. Here, we follow the most common choice in the literature (e.g., [20,21,46]) by applying the use of the gross domestic product (GDP). GDP represents the level of human activities in a period. Furthermore, by including the added value contribution by the capital factor, GDP also captures the level of the stock of assets present in the area.

GDP is not readily available at the scales of Census blocks and for quarterly time intervals. We, thus, compute county level quarterly GDP values as in [47] for each 2-digit NAICS industry. The counties considered are the ones intersecting CRB as in Figure 2. We divide the GDP industry-level value by the number of jobs in the corresponding industry obtaining a GDP value per job of each industry in each county. We obtained the block level total GDP by multiplying those numbers with job information at the 2-digit NAICS that is reported in the workspace characteristic data (WAC) of Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES) database administered by the U.S. Census. The time intervals in the period 1997–2017 that are not

provided by LODES have been estimated by linearly extrapolating the jobs series for each block and industry.

When running future scenarios, we keep exposure constant and equal to the values observed in each quarter and block of the last year of the calibration period, i.e., 2017. For instance, the exposure of block b in the second quarter of 2078 will be as the GDP in the same block in the second quarter of 2017. In this manner and by using exposure values at 2017 prices as obtained by correcting current values with the consumer price index (CPI) released by the Bureau of Labor Statistics (BLS), we keep prices constant for both model calibration and predictions.

2.1.3. Damage

We consider only the direct economic losses (DEL) that have been observed for events recorded in the Storms Events Database published by the National Centers for Environmental Information (NCEI) of the U.S. National Oceanic and Atmospheric Administration (NOAA).

We first split each event in the Storm Events Database to blocks and quarters as follows. From the temporal perspective, we proportionally distribute the events according to their start and end time to single or multiple quarters. From the spatial perspective, we use the latitude and longitude of extent of each event to split it evenly across the blocks intersected by the event. The value of damage distributed to blocks and quarters either proportionally or evenly is computed as the sum of damage to crops, properties, and people. Damage to properties and crops are used as reported in the original database while injuries and fatalities are quantified as it follows. Fatalities are transformed in monetary values through the value of statistical life (VSL) as computed nationally by the U.S. Department of Transportation (DOT) based on income losses with elasticity as in [48] and personal income data published by the U.S. Bureau of Economic Analysis (BEA). Injuries are converted to monetary values by referring to moderate severity injuries that are valued at 4.7% of VSL. Damage values are reported at 2017 prices.

2.2. Model calibration

We use the dataset described above for the period 1997–2017 to calibrate the adaptation model. In particular, we consider only blocks where the considered exposure has been positive sometime in the period. Summary statistics of the calibration sample are reported in Table 2.

Table 2. Summary statistics for calibration sample considered. Following recent empirical modeling approaches [19,20,21,49] to the adaptation to climate disasters and extremes, we calibrate the adaptation model by using count variables regression approaches and accounting for statistical over dispersion in the data. In particular, we fit a zero-inflated negative binomial (ZINB) model to overdispersed count data with excess zero counts. The ZINB model assumes that the excess zero counts come from a logit model and the remaining counts come from a negative binomial model. The probability distribution of the ZINB random variable can be expressed as follows:

$y_i \sim \{0NB(\mu_i, \theta)\}$ with probability p_i with probability $1-p_i$

(2)

where y_i is our dependent count variable that is the non-negative

value in dollars of damage in a block, and θ is the dispersion parameter (the estimate of this parameter points to over-dispersion and, hence, here a negative binomial model is more appropriate than a Poisson model).

Hence, the regression models, without reporting error terms, are a negative binomial regression model for positive values of the count variable damage:

$$\log(\mu_i) = \beta_0 + \beta_1 H_i + \beta_2 X_i + \beta_3 E[H]_i,$$

(3)

and a logit model to capture the zero inflation:

$$\text{logit}(p_i) = \gamma_0 + \gamma_1 H_i + \gamma_2 X_i + \gamma_3 E[H]_i,$$

(4)

where we consider the same covariates in both models (i.e., hazard-H, exposure-X, and hazard expectation-E[H]).

Estimates for β

, γ , and θ are obtained by maximizing the log-likelihood function.

$$L(\beta, \gamma, \theta | y, X) = \sum_{i=1}^n \ln \log f(y_i | X_i, \beta, \gamma, \theta).$$

(5)

We first select the hazard ESM that better fits the data. In fact, we run the estimation three times with different values of H and E[H]

, and in each run, we use the three different H presented above (i.e., A, B, and C as in Table 2) and their corresponding moving average over the previous 20 years as E[H] (i.e., for each block b at time t, we have $E[H]_{b,t} = \sum_{l=1}^L H_{b,t-l}$, where $L=80$ are the quarters in the previous 20 years). We then select the joint hazard prediction between A, B, C, and the average values obtained from linearly composing three ESMs and all the possible couples of them with respect to H and E[H]

that maximizes the log likelihood value computed as in Equation (5).

The selected ESM is B, i.e., the IPSL-CM5A-LR.

Second, we fit the model by obtaining the estimates reported in Table 3. Values reported in Table 3 should be interpreted as incident rate ratios (IRR) and odds ratios and the constant reports the baseline IRR and odds ratio. Increases in probabilities reported for covariates in the binomial model refer to increases in the positive value of the dependent variable (i.e., damage). Similar increases in the logit model refer to the zero-inflation model and, thus, refer to increases in probabilities to obtain zero-value damage. Results not reported here confirmed the need for modeling zero inflation and for choosing a ZINB model over a zero inflated Poisson (ZIP) model.

Table 3. Adaptation model calibration: zero-inflated negative binomial estimates of flooding damage (USD) in the CRB 1997–2017 (monetary values in 2017 prices).

Odds ratios and IRRs in Table 3 and the coefficients derived largely confirm theoretical model predictions. First, in the inflation model, the probability of having no damage is negatively correlated with H and X and positively with E[H]

. Second, in the negative binomial model, E[H]

is negatively correlated with damage. In this latter model, the

significant coefficient of X has the opposite direction of what is predicted by the theory, and this could be interpreted from a two-fold perspective. First, the total impact is on the right direction since the inflation model coefficient is larger. Second, this could be explained as further evidence in favor of larger preferences for safety by wealthier communities, as found in many empirical works (e.g., the inverted u-shape damage function in [20]).

3. Results and Discussion

3.1. Floodings as Concurrent Extremes

Copulas provide a means to represent future concurrency of the types of events that can result in flooding, such as maximum soil moisture, runoff, and precipitation, as applied in this work.

However, copulas as a tool are difficult to apply when attempting to link results to actual occurrences of events. For instance, direct comparisons to flooding values in the CRB are next to impossible given the rare nature of events, and the variability of these events within the historical record. Hydrologic models, such as VIC, when coupled with ESMs, provide us a tool to examine future changes, but they are less than ideal to consider current and near-term events; thus, this is a limitation of the work presented herein.

Despite the uncertainties associated with the use of copulas, we can use copulas to detect flooding occurrences in historical and future records and discuss the changes and variability over time. Average future values of flooding hazards until 2100 are reported for the IPSL-CM5A-LR ESM, which was chosen as the best fit to historical data in calibration (Figure 6). This ESM is also one of the driest scenarios for CRB; thus, our results represent a moderate extrapolation of flooding risks under future climate change in the CRB.

Figure 6. Quarterly level of flooding hazard as the predicted average levels in CRB Census blocks until 2100. Predictions based on IPSL-CM5A-LR ESM.

Counts of flooding instances across the entire CRB (grid cells) for each ESM are provided in the Supplemental Materials (Figure S1) and coincide with the average level of natural hazards shown here.

Flooding instances are low throughout the 2020–2040s, and then it rises after 2040 to a peak in 2060 to 2080s and are generally lower after the 2080s, coinciding with a decrease in snow in the basin (e.g., [50]). IPSL-CM5A-LR ESM, as the driest of all ESMs, has the lowest number of flooding occurrences (maximum value of 123), with 159 months with flooding events occurring in them over the entire record. MIROC-ESM, the hot/wet ESM, has the highest number of flooding occurrences (256), while the warm/wet ESM, GFDL-ESM2G, is in between the two in terms of the number of occurrences (185). On the other hand, both GFDL-ESM2G and MIROC-ESM contain the same number of flooding months, 179, only 20 more than IPSL-CM5A-LR. GFDL-ESM2G tends to have more flooding occurrences after the 2060s, and even IPSL-CM5A-LR appears to have a final spike in the record right before the end of the century.

Results about floodings as concurrent extremes point out that, overall, flooding is highly variable, likely owing to the competing nature of increasing precipitation and decreasing snow, with steadily increasing temperatures, under future climate scenarios. The results are consistent with main outcomes of contributions using

similar approaches applied to different regions and based on different ESMs [51,52].

3.2. Adaptation Model

Modeling adaptation as an investment problem means identifying three main variables of interest. The first one is the natural hazard H , which is to say a quantification of intensity and frequency of extremes. The second one is exposure X , which is a single monetary value that represents all human assets and activities at stake. Finally, there is damage D , which is the monetary value of human assets and activities that are lost because of the occurrence of natural hazards.

The adaptation model assumes that D is a positive function of both H and X and that the difference between a more adapted system and a less adapted one is that facing the same levels of X and H where the former has a lower D than the latter. Furthermore, while economic theory tells us that the optimal level of investment in adaptation is the one where the adaptation marginal benefit equals its marginal cost [53], such theoretical predictions are currently difficult to use for local communities and policies since we neither have means to effectively measure the quantity of adaptation nor its cost function (benefits could be measured as decreases on predicted D levels, which is to say as avoided losses).

Given the difficult applicability of theoretical results, here we improve the adaptation model by adding a fourth variable $E[H]$

, which is the expected level of natural hazards. Expected hazard drives the assessment of future potential damage and, hence, of the future benefits of adaptation investments. Expected hazard should, thus, be negatively correlated with D . Our resulting model is, thus, $D=f(H+X+E[H]-)$. The analysis in [21] demonstrated how, consistently with the adaptation model based on investment choices, expectations about future exposure $E[X]$

are also negatively correlated with D . Here, another type of expectation that could be used for policy making and that supports the positive role and impacts of scientific research on climate change is investigated.

Results of our model calibration for the CRB over the period 1997–2017 (see details in Section 2) are largely consistent with theoretical predictions. Using these results, we can examine the empirically calibrated model by predicting D over H and considering different levels of $E[H]$

and X as observed on average across CRB blocks in 2017. Predicted quarterly levels of damage are presented in Figure 7 for increasing levels of H on a logarithmic scale and up to the maximum level observed in a single CRB block (i.e., $H=57$).

Figure 7. Adaptation model results. Predicted damage over levels of natural hazard (flooding) and scenarios of hazard expectations (average 2017 damage D and hazard $H = 1$) with constant exposure (exposure $X =$ average 2017 value).

The three different levels of $E[H]$

tells us different things. One scenario, represented as a green line, represents the expectation of future H as observed on average in 2017 (i.e., $H=0.6$). A second scenario (blue line) is to have

expectations as those calibrated in 2017 ($E[H]=0.365$), which is the quarterly mean across blocks of the moving average of H over the previous 20 years. Finally, the red line illustrates the results for hazard expectations equal to the average H experienced in the last two decades of the last century ($E[H]=0.164$

). The three scenarios represent three hypotheses of decreasing speed in the process by which local communities adapt their expectations about the future to a changing climate and associated changing levels of natural hazard.

Results highlight that communities that are capable of updating their expectations of natural hazards to levels recently experienced can adapt more, up to the point where they can tolerate much more natural hazards than others. For instance, this is confirmed by looking at the very similar D obtained when comparing the red-line scenario (i.e., $E[H]=H$

as in 1980–2000) for $H=20$ and the green-line scenario (i.e., $E[H]=H$ as in 2017) for $H=50$

. By keeping constant the human assets and activities at stake, future damage levels can be up to 12-times what was observed in 2017, confirming to a large extent results obtained with other approaches [54,55,56]. However, we can now add that this is true only if more severe hazards are not expected and that actually expectations have the potential to offset the adverse impacts of a changing climate through adaptation.

3.3. Scenarios until the End of the Century

We use predictions about flooding hazards in CRB until the end of this century to study three scenarios of $E[H]$

designed above to investigate the impact on adaptation of quickly updating hazard expectations. In particular, we consider a scenario of a constant level of $E[H]$ equal to the average H experienced by each block in the 1980–2000 period. Second, we investigate a dynamic scenario where $E[H]$ is updated in each block at the beginning of a new decade by using the average value of the experienced H in the two previous decades. Finally, we introduce a scenario in which the hazard expectation is also partially influenced by scientific predictions of natural hazard levels. In this scenario, at the beginning of each decade, $E[H]$

in each block is updated to be equal to the average actual H in the last decade and the one correctly predicted for the current decade. We run the adaptation model in each CRB block quarterly obtaining results as in Figure 8. The figure confirms most of the dynamics suggested by H predictions (see also Figure 6) and points out that up-to-date $E[H]$

lower predicted damage even in periods of relatively low levels of natural hazard (see, for instance, the difference between the blue line and the others in the years before 2040).

Figure 8. Quarterly flooding damage until 2099 in CRB. Predicted average levels in census blocks by hazard expectation scenario. Furthermore, we remove seasonal dynamics by aggregating predictions in annual levels and we present them in Figure 9 as the decadal mean and 95% confidence intervals to compare future scenarios with predicted and observed levels in the adaptation model calibration

period 1997–2017 (the 21-year long calibration period is reported at the beginning of the horizontal axis of the figure).

Figure 9. Annual flooding damage from 1997 to 2099 in the CRB.

Observed and predicted damage in census blocks. Average levels and 95% confidence intervals computed over adaptation model calibration period, decades, and hazard expectation scenarios.

As noted previously, updating hazard expectations continuously improves adaptation (i.e., reduces predicted damage) even when hazard levels are relatively low (see for instance decades starting in 2020, 2030, 2060, and 2080). Furthermore, updating expectation by including also predictions for the hazard in the current period strongly improves adaptation in particular in periods when natural hazards are at peak levels. The generic adaption model (and with other approaches such as those reported in [54,55,56]) results are confirmed by fine-scale predictions of the CRB for the remainder of this century. In terms of the size of what can be gained by using up-to-date and even forward-looking $E[H]$

, values such as those in the 2070–2079 decade point out that the decrease in predicted damage can be large and, on average, even half of the total experienced in 1997–2017 and in some years even four or five times that value.

4. Conclusions

We show here how to model the endogenous process of human adaptation to extremes as an investment decision. We apply the approach to flooding in the U.S. Colorado River basin, developing and adapting the data and approaches needed, such as the one considering the concurrent distribution of several variables for the quantification and prediction of climate extremes. Our approach is independent of the geographical area and the extreme event considered in this work. The method can be replicated on other areas in the world and focuses on different kinds of extremes (e.g., droughts, wildfires, snow and ice storms, tornadoes, etc.).

Results from both the theoretical (but empirically calibrated) model and from a fine scale model studying future predicted extremes up to 2100 show that hazard expectation has the potential to offset adverse changes in damages due to a changing climate [22]. We use the driest ESM (i.e., IPSL-CM5A-LR); thus, our results are robust even under the lowest future flooding projection. Changes in future damage due to hazard expectations are significant even though changes in extremes are conservative. The findings presented in [21] show how expectations about future exposure can shape adaptation. Expectations about the future (either about future exposure or future hazard) are, thus, a predominant factor in adaptation choices by human systems, and these choices can be studied as investments aimed at managing future risks.

Our results also indicate that the speed by which local communities adapt their expectations about the future to a changing climate and associated changing levels of natural hazard is key in managing future risks. Recent empirical studies in the same area [57,58] have pointed out how the speed of expectation update is strongly correlated with historical experience. Communities that have faced the most adversities most recently are the quickest to update their beliefs about a different future. We can conclude that particular attention should be spent on communities that, by chance, have not

experienced major recent extremes or that have experienced a recent sharp increase in exposure because they are not only the ones most likely to have erroneous expectations about future natural risks but also the ones that are the slowest in updating them.

Modeling human systems adaptation to extremes as an investment choice effectively captures main variables and dynamics at work and allows us to quantify the impact of hazard expectations [19,20,21]. This variable influences today's investments in adaptation by defining their future benefits, and as such plays a critical role in adaptation that calls for quantification and accurate modeling. The research supports the use of expectations as key policy lever for adaptation. Effective adaptation strategies could be based on assessing losses potentially caused by future natural hazards. Local communities would, thus, be allowed to fully understand the risk posed by climate change. Updating community beliefs is a complex learning process that can be supported by public policies [59,60,61].

In this context, scientific credibility is also extremely important because risk assessment for the future must be trustworthy in order to be accepted and used by society. In fact, scientific predictions are often difficult to be effectively incorporated by public policies in particular when they are derived from advanced but complicated models, which are difficult to be explained and trusted [62]. Science reputation is key for acceptance [63,64,65] and for protecting communities through the adaptation mechanism investigated here. In order to establish and maintain scientific credibility, society has to invest in particular to advanced prediction accuracy. Our results, thus, suggest a positive benefit of public investments in climate impacts studies on human systems. Good science makes human systems adaptation possible and effective.