

Ratemaking and Product Information

Catastrophe Insurance Pricing: An Adaptive Robust Optimization Approach

Cynthia Zeng^{1a}, Dimitris Bertsimas^{2b}

¹ Meteorologica Technologies, ² Massachusetts Institute of Technology

Keywords: Catastrophe Insurance, Optimization

<https://doi.org/10.66573/001c.167995>

Variance

Vol. 19, 2026

The escalating frequency and severity of natural disasters, exacerbated by climate change, underscores the critical role of insurance in facilitating recovery and promoting investments in risk reduction. The paper introduces a novel adaptive robust optimization (ARO) framework tailored for the calculation of catastrophe insurance premiums, with a case study applied to the US National Flood Insurance Program. To the best of our knowledge, it is the first time an ARO approach has been applied to disaster insurance pricing. Our methodology is designed to protect against both historical and emerging risks, the latter predicted by machine learning models, thus directly incorporating amplified risks induced by climate change. Using US flood insurance data as a case study, optimization models demonstrate effectiveness in covering losses and produce surpluses, with a smooth balance transition through parameter fine-tuning. Among tested optimization models, results show that ARO models with conservative parameter values achieve a low number of insolvent states with the least insurance premium charged. Overall, optimization frameworks offer versatility and generalizability, making them adaptable to a variety of natural disaster scenarios, such as wildfires and droughts, among others. This work not only advances the field of insurance premium modeling but also serves as a vital tool for policymakers and stakeholders in building resilience to the growing risks of natural catastrophes.

1. INTRODUCTION

Catastrophic events—ranging from natural disasters to human-made crises—can cause severe damage to individuals, businesses, and entire communities. In recent years, global climate change has exacerbated climate variability and extreme weather events, leading to more frequent and costly natural disasters worldwide (Van Aalst 2006). As [Figure 1](#) shows, the frequency of disasters has risen dramatically over recent decades. The heightened risks of catastrophic events, including tropical cyclones, droughts, floods, and heatwaves, pose unprecedented challenges to society's resilience and adaptability. These challenges have made disaster risk mitigation and climate change adaptation strategies increasingly urgent (Mercer 2010).

The Intergovernmental Panel on Climate Change has underscored the necessity of financial instruments to manage disaster risk and support climate adaptation efforts (Linnerooth-Bayer and Hochrainer-Stigler 2015). Developing innovative tools for managing financial exposure to weather risks is crucial. The insurance sector has emerged as a pivotal player in protecting human, economic, and natural systems from climate-related vulnerabilities (Mills 2007; Keucheyan 2018). Within this context, catastrophe insurance serves as a vital risk management mechanism, not only providing financial support for post-disaster recovery but also creating incentives for proactive investment in mitigation efforts.

Catastrophe insurance, also known as disaster insurance, focuses on large-scale, low-frequency events that can cause

a Cynthia Zeng is the founder of Meteorologica Technologies, a weather AI company. She received her PhD from MIT, where her dissertation examined multimodal machine learning methods for climate adaptation. She subsequently held a tenure-track appointment as assistant professor at NYU Stern School of Business. Her research has been published in peer-reviewed journals, including *npj Digital Medicine* and *Weather and Forecasting*. Address for correspondence: cynthia@meteorologica.com.

b Dimitris Bertsimas is the vice provost for Open Learning, associate dean of Online Education and Artificial Intelligence, Boeing Professor of Operations Research, and faculty director of the Master of Business Analytics program at MIT, where he has been a faculty member since 1988. His research concerns optimization, machine learning, and applied probability, with applications in healthcare, finance, operations management, and transportation. He is the author of more than 350 scientific papers and eight graduate-level textbooks.

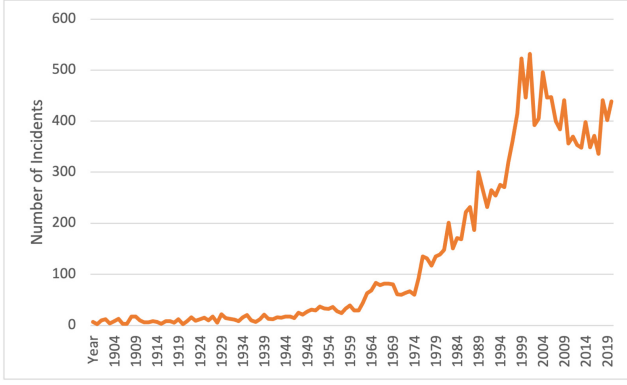


Figure 1. Number of major disasters globally since 1900, maintained by the EM-DAT database (“EM-DAT The International Disaster Dataset” 2021). A disaster is defined as an event that overwhelms local capacity, necessitating a request to the national or international level for external assistance. Types of disaster include flood, storm, earthquake, drought, landslide, extreme temperature, wildfire, volcanic activity, mass movement (dry), glacial lake outburst, and fog, among others.

widespread damage. It typically covers disasters such as hurricanes, earthquakes, floods, terrorist acts, and pandemics. The rarity of such catastrophic events complicates the insurance process, as traditional actuarial methods fall short, often due to a lack of comprehensive historical data. Compounding that challenge, climate change is leading to more regular and destructive climate-related catastrophes, which traditional reliance on historical data alone tends to underestimate. In contrast to conventional insurance policies that spread risks across insured individuals, catastrophe insurance confronts a temporal problem of matching the regular influx of annual premiums with the irregular and unforeseeable distribution of payouts for losses.

In the United States, catastrophe insurance has historically been managed predominantly through national programs. The National Flood Insurance Program (NFIP), for example, is the principal provider of flood insurance in the country, covering more than 95% of the underwriting risks (Michel-Kerjan 2010). However, the NFIP’s actuarial effectiveness has been the subject of scrutiny, with the program operating at a significant deficit—US\$19 billion as of 2023—highlighting the need for reform in the structuring of such insurance schemes (Horn and Brown 2017). The entry of private insurers into the catastrophe coverage market has been deemed crucial, yet the inherent complexities associated with rare events have led to minimal participation from private entities (Michel-Kerjan and Kunreuther 2011). Indeed, many private insurers are withdrawing from regions deemed “uninsurable” due to escalating risks from climate change (Flavelle et al. 2023). Regulators and insurers are thus particularly concerned about insolvency risks in catastrophe insurance markets, as extreme events can trigger large-scale, correlated losses that threaten insurers’ financial stability. These solvency concerns have led to stringent capital requirements and risk transfer mecha-

nisms, such as reinsurance and catastrophe bonds, to enhance the resilience of insurance providers (Charpentier and Le Maux 2014).

In this paper, we present an adaptive robust optimization (ARO) framework for catastrophe insurance premium pricing designed to protect against uncertain losses. To the best of our knowledge, this is the first work using an ARO approach to set disaster insurance premiums. We develop the framework and implement it to flood insurance using NFIP data. The main contributions are threefold:

- We present an ARO framework designed for the pricing of catastrophe insurance premiums. The novel framework integrates two key objectives: minimizing insolvency risk and ensuring capital efficiency. Specifically, we develop an uncertainty set that captures the volatile loss profile based on historical loss distributions, enhancing insurers’ resilience against catastrophic events.
- We apply our ARO framework to US flood insurance using data from the NFIP from 1975 to 2022. We parameterize optimization models using 1975–2011 data as training data and evaluate model performance using 2013–2022 data as out-of-sample testing data. Optimization models demonstrate capabilities in effectively covering losses while offering adaptability to policymakers’ risk tolerance and level of conservatism. In particular, we recommend that policymakers choose an ARO model, with conservative parameter values, to achieve superior performance in both effectiveness and efficiency in covering losses.
- We highlight the adaptability and generalizability of our framework, suggesting the potential application of an ARO approach to pricing a wide range of catastrophic events, such as wildfires, droughts, and extreme weather events.

The structure of the paper is as follows. In Section 2, we review the relevant literature. In Section 3, we introduce the problem setup and present both the robust optimization (RO) and the ARO frameworks. We derive their respective robust counterparts in forms that can be efficiently solved using standard optimization solvers. In Section 4, we demonstrate the application of our framework through a case study on flood insurance in the United States. We explain model construction and model parameter estimation. In Section 5, we discuss numerical results of the case study from our models against two baselines: historical NFIP premiums and a cumulative moving average (CMA) scheme. Finally, we conclude in Section 6.

2. LITERATURE REVIEW

The problem studied in this paper pertains to three key areas: (1) catastrophe insurance pricing, (2) catastrophe modeling, and (3) disaster relief planning. Our methodology builds on research in adaptive and robust optimization.

2.1. CATASTROPHE INSURANCE PRICING

The cost of catastrophe insurance comprises three key components: average annual loss (AAL), risk load, and expense load (Grossi 2005). AAL represents the total expected loss from all covered events within a year. The expense load accounts for administrative costs, including taxes, commissions, and loss adjustment expenses, typically set as a fixed percentage (e.g., 30%) of the combined AAL and risk load. Unlike conventional property and casualty insurance (e.g., auto insurance), where losses remain relatively stable across large pools and AAL constitutes the bulk of the premium, catastrophe losses are characterized by extreme volatility. Insurers need to allocate substantial additional capital beyond the expected loss to ensure that they can absorb rare shocks, and the risk load reflects such a cost premium.

To mitigate insolvency risks, regulators require insurers to adhere to capital adequacy principles, often through exceedance probability thresholds to ensure that the probability of total losses surpassing available capital remains acceptably low (Olivieri and Pitacco 2015). The required capital varies based on insurer-specific strategies and regulatory mandates (Jaffee and Russell 1997). For example, the Florida Office of Insurance Regulation stress-tests hurricane insurers to ensure that the probability of insolvency does not exceed 1% (Nicholson et al. 2018). Instead of decomposing risks into AAL and risk load, our approach adopts an optimization framework, directly modeling the cost required to prevent insolvency risk by constructing uncertainty sets around expected losses. This allows for a more flexible and robust risk assessment, accounting for volatility without relying solely on predefined risk load assumptions.

2.2. CATASTROPHE MODELING

Catastrophe modeling, or CAT modeling, is a computational framework that integrates statistical, geophysical, and economic data to estimate the likelihood and financial impacts of extreme events such as hurricanes, earthquakes, and floods. Widely used by insurers and policymakers, CAT models assess disaster probabilities and estimate physical and financial losses (Grossi 2005). Modern CAT models consist of three key modules: hazard modeling, vulnerability assessment, and exposure modeling (Mitchell-Wallace et al. 2017). Hazard modeling simulates event characteristics (e.g., wind speed, earthquake magnitude) using statistical and physics-based methods, with machine learning improving predictive accuracy (Rueden et al. 2020; Chen and Zhang 2022). Vulnerability assessment quantifies the relationship between hazard intensity and damage, while exposure modeling links predicted damages to affected assets and populations through high-resolution geospatial and socioeconomic datasets. The market is dominated by proprietary models from AIR Worldwide, RMS, and EQECAT, alongside open-source alternatives like the Federal Emergency Management Agency's (FEMA's) HAZUS (Schneider and Schauer 2006).

Despite their widespread use, CAT models face challenges such as limited historical data, complexities in scenario simulations, and the underestimation of risks, particularly under climate change (Kunreuther et al. 2013; Dlugolecki et al. 2009). Emerging hybrid approaches that combine machine learning with physics-based models offer improved transparency, robustness, and adaptability (Zahura et al. 2020). This paper contributes to the literature by introducing an optimization-based framework that directly addresses uncertainties in losses and climate risks, enhancing tail-event risk assessment and improving decision-making for risk management.

2.3. DISASTER RELIEF PLANNING

Many studies have focused on the problem of general resource allocations to different programs or regions under a given budget constraint (Wang et al. 2023; Alem et al. 2016; Salmerón and Apte 2010). Yang et al. (2021) discuss fund allocation for flash flood reduction, and Liu et al. (2022) further incorporate the use of insurance premiums as a source of funding. As highlighted by Kunreuther et al. (2013), the critical issue in insurance policy lies in the need for decision-making robustness in the face of climate change's uncertainties. There is also a growing body of work on using robust optimization in disaster relief management to deal with uncertainties (Ben-Tal et al. 2011; Zokaee et al. 2016). Few studies discuss the use of optimization for catastrophe insurance pricing directly. Ermolieva et al. (2017) apply stochastic optimization to the Dutch flood insurance scheme. Our study fills in the gap by proposing a robust optimization framework by modeling uncertain losses to address the increasing unpredictability of weather events driven by climate change.

3. OPTIMIZATION FRAMEWORK FOR CATASTROPHE INSURANCE

In this section, we introduce a realistic problem setup from the perspective of an insurer, where annual premium rates must be determined under significant uncertainty about future losses. Losses due to natural disasters and other extreme events are inherently uncertain, often exhibiting heavy tails, temporal variability, and limited data availability. To account for such uncertainty in a systematic and risk-averse manner, we adopt an RO framework. RO provides a natural approach for modeling uncertainty by optimizing decisions against the worst-case outcomes within a predefined uncertainty set, which is defined using historical loss distribution. We then present the robust counterpart of the RO formulation, which can be solved efficiently using standard optimization solvers. To further enhance flexibility, we introduce an ARO framework, which allows rate-setting decisions in future years to adapt to realized losses, reflecting the evolving information available to insurers over time. Finally, we derive the robust counterpart of the ARO formulation.

3.1. PROBLEM SETUP

In this study, we consider setting insurance premiums for N locations for the insurance period of T years, denoting premiums as variables $p_{i,t}$, where $i \in [N], t \in [T]$. Following industry practice, we assume that each location represents the smallest geographical unit where catastrophe risk has been assessed and that a standard premium is charged. We are given the data of historical losses for each location for each year in the past T_0 years, which we denote with $l'_{i,t}$, $i = 1, \dots, N$, $t = 1, \dots, T_0$. We assume that the future losses for each location i in the insurance period of T years are given by $l_{i,t}$, $i = 1, \dots, N$, $t = 1, \dots, T$. Future losses $l_{i,t}$ for location i over the insurance period T are inherently uncertain. However, to first establish our formulation framework, we assume knowledge of $l_{i,t}$ and later introduce an RO approach to systematically model this uncertainty.

The objective is to determine the minimum premium that will ensure financial stability while offering the most competitive rate in the market. This involves optimizing capital efficiency, specifically, covering insolvency risks with the least amount of required capital. By minimizing capital requirements, insurers are better positioned to lower premiums and remain competitive. To account for consumer behavior, where higher premiums lead to decreased demand, we introduce a demand-damping function $f: \mathbb{R} \rightarrow [0, 1]$, which is a monotonically decreasing function representing the decline in demand as premiums increase. Therefore, the overall objective function of the optimization problem is as follows:

$$\min_{p_{i,t}} \sum_{i=1}^N \sum_{t=1}^T f(p_{i,t}) * p_{i,t}. \quad (1)$$

To model the damping behavior of demand as price increases, we adopt a piecewise linear function. This choice ensures that the resulting objective function remains quadratic in the premium variable p , preserving the tractability for most optimization solvers to find a solution efficiently. Moreover, the damping rate can be flexibly controlled through a tunable hyperparameter, allowing us to accurately reflect the sensitivity of demand to price changes.

Next, to protect insurers against insolvency risk, we require the premium price to cover projected losses with an additional buffer amount, denoted by δ , which is a fixed threshold for each location. To ensure financial stability, then, premiums must exceed projected losses by the buffer amount, leading to the following constraint:

$$\sum_{t=1}^T f(p_{i,t}) * p_{i,t} - \sum_{t=1}^T f(p_{i,t}) * l_{i,t} \geq \delta, \quad i \in [N]. \quad (2)$$

Next, we impose a constraint to require premiums collected over consecutive years to vary slowly. Insurance regulators typically impose limits on year-over-year premium increases to prevent market disruptions and financial strain on policyholders. Thus, we enforce the following constraint to require premiums collected in subsequent years to be bounded by an absolute value θ :

$$|p_{i,t} - p_{i,t-1}| \leq \theta, \quad i \in [N], \quad t \in [T]. \quad (3)$$

Finally, $p_{i,t}$ should be nonnegative, for all locations over all periods, to ensure feasibility in all scenarios:

$$p_{i,t} \in \mathbb{R}^+, \quad i \in [N], \quad t \in [T]. \quad (4)$$

3.2. ROBUST OPTIMIZATION FORMULATION

In our initial problem setup, we assume that projected future losses are known. However, in reality, future losses $l_{i,t}$ for location i over the insurance period T are inherently uncertain and highly volatile, particularly in the context of catastrophe risk. To ensure that insurance premiums remain resilient across different possible loss scenarios, we adopt an RO framework to systematically account for uncertainty.

RO is a well-established methodology for handling optimization problems with uncertain data (Bertsimas et al. 2011). Unlike traditional stochastic approaches that rely on probability distributions, RO optimizes against an uncertainty set, making it particularly suited for extreme events and financial risk management. Given the unpredictability of catastrophic losses, RO naturally lends itself to insurance premium modeling, allowing insurers to mitigate insolvency risks while maintaining capital efficiency. Widely applied in operations research, engineering, and finance, RO provides a structured, scenario-resilient framework that balances risk and financial stability, ensuring that premiums remain robust across uncertain loss scenarios.

We model the uncertain variables $\{l_{i,t}\}_{i \in [N], t \in [T]}$ using uncertainty sets $\{\mathcal{U}_i\}_{i \in [N]}$ and impose the requirement that constraint (3) holds for all possible values within the uncertainty set. This ensures that premiums remain sufficient to cover potential losses under all scenarios captured by the uncertainty set. Thus we ensure that premiums are robustly set to mitigate insolvency risks to maintaining financial sustainability:

$$\begin{aligned} \sum_{t=1}^T f(p_{i,t}) * p_{i,t} - \sum_{t=1}^T f(p_{i,t}) * l_{i,t} &\geq \delta, \\ \forall (l_{i,t})_{t \in [T]} &\in \mathcal{U}_i, \\ i &\in [N]. \end{aligned} \quad (5)$$

To account for uncertainty in future flood losses, we construct an uncertainty set that captures potential deviations from historical data. We assume that future flood losses at each location follow the same distribution as past losses, and we leverage the central limit theorem (CLT) to define a robust uncertainty set. For each location i , we assume that the future losses $L_{i,t}$ are independent and identically distributed (i.i.d.) with a mean of $\bar{l}_i$ and a standard deviation of $\bar{\sigma}_i$, where $\bar{l}_i$ and $\bar{\sigma}_i$ are the mean and the standard deviation for location i estimated using historical data. We assume that the uncertain quantities $L_{i,t}$ take values such that

$$\left| \sum_{t=1}^T L_{i,t} - T \cdot \bar{l}_i \right| \leq \gamma \cdot \sigma_i \sqrt{T}, \quad (6)$$

where γ is a parameter controlling the conservatism of the uncertainty set, determining how much future losses can deviate from the historical distribution. Thus, the uncertainty set $\mathcal{U}_i$ captures all possible future loss scenarios for location i where cumulative losses remain within a bounded deviation from the historical mean, scaled by γ

$$\mathcal{U}_i = \left\{ (l_{i,1}, \dots, l_{i,T}) : \frac{|\sum_{t=1}^T l_{i,t} - T \cdot \bar{l}_i|}{\sigma_i \sqrt{T}} \leq \gamma \right\}, \quad (7)$$

where $\bar{l}_i, \sigma_i$ can be computed for each location i using historical data. Increasing γ expands the uncertainty set, making the optimization model more conservative. This results in higher premiums to ensure sufficient coverage under worst-case loss scenarios. Since flood risk varies across locations, we construct a separate uncertainty set for each location using its respective historical mean and standard deviation. This ensures that our model captures regional differences in flood risk while maintaining robustness across all locations.

Therefore, the overall RO formulation of the problem is as follows:

$$\begin{aligned} \min_{p_{i,t}} \sum_{i=1}^N \sum_{t=1}^T f(p_{i,t}) * p_{i,t} \\ \sum_{t=1}^T f(p_{i,t}) * p_{i,t} - \sum_{t=1}^T f(p_{i,t}) * l_{i,t} \geq \delta, \quad (8) \\ \forall (l_{i,t})_{t \in T} \in \mathcal{U}_i, \quad i \in [N], \\ |p_{i,t} - p_{i,t-1}| \leq \theta, \quad i \in [N], t \in [T], \\ p_{i,t} \in \mathbb{R}^+, \quad i \in [N], t \in [T]. \end{aligned}$$

Proposition 3.1. The RO formulation is equivalent to

$$\begin{aligned} \min_{p_{i,t}} \sum_{i,t} f(p_{i,t}) * p_{i,t} \\ \sum_{t=1}^T f(p_{i,t}) * p_{i,t} - \frac{1}{T} \sum_{t=1}^T f(p_{i,t}) * L_i \geq \delta, \quad \forall i \in [N], \quad (9) \\ |p_{i,t} - p_{i,t-1}| \leq \theta, \quad \forall i \in [N], t \in [T], \\ \text{where } L_i = T \cdot \bar{l}_i + \gamma \cdot \sigma_i \sqrt{T}. \end{aligned}$$

3.3. ADAPTIVE ROBUST OPTIMIZATION FORMULATION

ARO extends the RO framework introduced in the previous section by allowing the decision variables to depend on realized uncertain quantities through affine decision rules, rather than being fixed in advance against the worst case. Using ARO techniques, we enable premium adjustments based on actual loss experiences. In particular, we let premiums depend on realized losses using affine decision rules, as proposed in Chapter 7 of Bertsimas and Hertog (2022). This approach not only refines premium pricing accuracy but also ensures a responsive and equitable insurance mechanism against the backdrop of unpredictable catastrophic events. We let premiums depend on loss from the previous time step as follows:

$$p_{i,t} = \begin{cases} \alpha_{i,1}, & \text{for } t = 1 \\ \alpha_{i,t} + \beta_{i,t} \cdot l_{i,t-1}, & \text{for } t = 2, \dots, T \end{cases} \quad (10)$$

where premium for location i at time period t is determined by a linear combination of parameters. Specifically, for the first time period, the premium is set to a base value $\alpha_{i,1}$; for subsequent periods, the premium is adjusted based on the loss $l_{i,t-1}$ experienced in the previous period, with $\alpha_{i,t}$ and $\beta_{i,t}$ new variables to be optimized over.

Throughout this section, we drop the demand-damping function $f(p)$ from the ARO formulation. This simplification is a technical necessity: Retaining $f(p)$ together with the affine decision rules, in which $p_{i,t}$ depends on the uncertain loss $l_{i,t-1}$, would render the problem non-convex and preclude derivation of a tractable robust counterpart. As shown in the sensitivity analysis in Section 5.1, the choice of demand-damping rate has a negligible effect on

optimization outcomes compared to gamma, further justifying this modeling choice. The ARO formulation is given as follows:

$$\begin{aligned} \min_{p_{i,t}} \sum_{i=1}^N \sum_{t=1}^T p_{i,t} \\ \sum_{t=1}^T p_{i,t} - \sum_{t=1}^T l_{i,t} \geq \delta, \quad \forall (l_{i,t})_{t \in T} \in \mathcal{U}_i, \quad i \in [N], \\ |p_{i,t} - p_{i,t-1}| \leq \theta, \quad i \in [N], t \in [T], \\ p_{i,t} \in \mathbb{R}^+, \quad i \in [N], t \in [T]. \end{aligned}$$

Proposition 3.2. The ARO formulation is equivalent to

$$\begin{aligned} \min_{\alpha_{i,t}, \beta_{i,t}} \Omega \\ \text{s.t.} \sum_{t=1}^T \alpha_{i,t} + \min_{s_1^1, s_2^1} \sum_{j=1}^2 c_j s_j \leq \Omega, \quad (12) \end{aligned}$$

$$\begin{aligned} s_1^1 - s_2^1 \geq \beta_{i,t}, \quad \forall t = 2, \dots, T, \\ s_1^1 - s_2^1 \geq 0, \\ s_1^1, s_2^1 \geq 0. \\ - \sum_{t=1}^T \alpha_{i,t} + \sum_j c_j s_j^2 \leq -\delta, \end{aligned}$$

$$\begin{aligned} s_1^2 - s_2^2 \geq 1 - \beta_{i,t}, \quad \forall t = 2, \dots, T, \\ s_1^2 - s_2^2 \geq 1, \\ s_1^2, s_2^2 \geq 0. \end{aligned} \quad (13)$$

$$\begin{aligned} \alpha_{i,t} + \sum_j c_j s_j^{3,t} \leq 0, \\ s_1^{3,t} - s_2^{3,t} \geq -\beta_{i,t} \\ s_1^{3,t} - s_2^{3,t} \geq 0, \\ s_1^{3,t}, s_2^{3,t} \geq 0. \end{aligned} \quad (14)$$

$$\begin{aligned} |\alpha_{i,t} - \alpha_{i,t-1}| \leq \theta_1, \quad t = 2, \dots, T, \\ |\beta_{i,t} - \beta_{i,t-1}| \leq \theta_2, \quad t = 2, \dots, T. \end{aligned} \quad (15)$$

Since the uncertain variables appear linearly in constraints, we derive the robust counterpart for each constraint by introducing slack variables for each constraint, where each set of the equations in the robust counterpart corresponds to specific constraints where uncertain variables appear. Constraint (12) is the epigraph formulation for the objective function, constraint (13) corresponds to covering losses, constraint (14) corresponds to the positivity constraint, and constraint (15) corresponds to the slowly varying constraint. Detailed derivation can be found in the appendix.

4. CASE STUDY ON US NATIONAL FLOOD INSURANCE PROGRAM DATA

In this section, we demonstrate the application of our framework through a case study on flood insurance in the United States. We explain the model parameter estimation and the details of the training of machine learning risk predictions.

4.1. DATA

We used two redacted datasets from the NFIP containing claims and policies, respectively. Both datasets are created and maintained by FEMA. The claims transaction data provide details on NFIP claim transactions across from all US states (FEMA 2023). This dataset consists of 2,570,089 lines

of claim transactions dated from 1970. Given the limited data availability in the early years, we included data from 1975 to 2022.

For this study, we aggregate data into the state level on an annual basis, and we have used the following features: date (“dateOfLoss”), state, and claim amount (“amount-PaidOnBuildingClaim”). Additionally, data from “MP” (Northern Mariana Islands), “AS” (American Samoa), “GU” (Guam), and “DC” (District of Columbia) have been omitted due to their limited data records. As a result, our cleaned dataset encompasses information from 52 jurisdictions over 48 years, including 50 US states alongside two territories recognized as island states—the US Virgin Islands and Puerto Rico—enhancing the geographical breadth of our study.

The policy premium data contain 228,664 lines of data encompassing policies from 2009 to 2022 (FEMA 2023). We use the policy data as a benchmark to compare model performance between last 10 years of the testing period from 2013 to 2022. Similar to the claims data, we aggregated data into the state level on an annual basis, and we have used the following features: state (“propertyState”), date (“policyTerminationDate”), and premium (“totalInsurancePremiumOfThePolicy”).

In this work, we consider setting insurance premiums at the state level on an annual basis. Specifically, we focus on premium setting for the 10-year period from 2013 to 2022. Model hyperparameters are validated using historical data from 1975 to 2012, and the performance of the rate-setting schemes under RO and ARO is evaluated over the testing period from 2013 to 2022. The claims dataset spans 1975–2022, and the policy dataset spans 2009–2022. We use claims data from 1975 to 2012 to estimate optimization parameters (historical mean and standard deviation of losses), and we use both claims and policy data from 2013 to 2022 as the out-of-sample testing period for benchmarking RO and ARO against the historical (Hist) and CMA schemes.

4.2. OPTIMIZATION MODEL PARAMETER ESTIMATION

Recall from equation (29), to compute L_i we need to compute the historical mean and variance for each state. We use the training data from 1975 to 2012 to estimate optimization model parameters to avoid data spoilage in the testing dataset. We compute the historical mean and standard deviation for all states on an annual basis. Table 4 in the appendix exhibits the historical mean and variance for the top-10 most costly states.

In this work, we model demand sensitivity to insurance premiums through a piecewise linear demand function. We estimate the decline rate using historical data from several states. We include Figures 4 and 5 in the appendix showing the scatterplot of the number of policyholders in a year against the mean policy premium of that state at that year. Different states have different degrees of sensitivity to price, but in general we observe a downward trend of declining policyholder numbers as a function of increased price. For the illustrative purpose of this work, we do not

specify different sensitivity in different states but use the same demand-damping function across all states.

We use the following piecewise linear demand-damping function to model demand damping:

$$f(p) = \begin{cases} 1, & \text{if } p \leq P_0, \\ 1 - m \cdot (p - P_0), & \text{if } p > P_0 \text{ and } f(p) \geq c_{\min}, \\ c_{\min}, & \text{otherwise.} \end{cases} \quad (16)$$

where P_0 is the minimum premium at which demand damping starts to occur, and $c_{\min}$ is the minimum fraction of demand. In this work, we choose P_0 to be $\frac{1}{10} \cdot P_{\text{hist}}^{\max}$, a fraction of maximum historical premium ever charged. We choose $c_{\min}$ to be 0.2, representing that at least 20% of the total demand is preserved regardless of price. We experimented with different demand-damping rates: $m = 1/P_{\text{hist}}^{\max}$. We include Figure 6 in the appendix to illustrate several choices of the demand-damping curve.

5. RESULTS

We implement two robust optimization models: RO and ARO. To evaluate their effectiveness, we compare the results against two baseline policy premiums:

- Historical premiums: The premiums previously charged in each state, referred to as “Hist”.
- Cumulative moving average (CMA) loss: The average loss up to a given year, denoted as “CMA”. The CMA premium is computed as follows:

$$p_{i,t}^{\text{CMA}} = \frac{1}{t} \sum_{t'=0}^t l_{i,t'}. \quad (17)$$

A summary of the implemented models is provided in Table 1.

We evaluate performance during the testing period, the last 10 years of available data in the NFIP dataset between 2013 to 2022. For the rest of the section, we choose θ to be 50,000 and δ to be 10,000. We also undertake a sensitivity analysis to assess the impact of our model’s parameter selections. Overall, we evaluate our model performance using the following two criteria:

- Effectiveness: To evaluate the effectiveness of models to cover losses at a state level, we count the number of insolvent states (i.e., the cumulative premium collected over the testing period does not cover cumulative claim losses).
- Efficiency: To ensure models are charging reasonable levels of premiums to cover losses, we evaluate the overall surplus (or deficit) level as well as the absolute deviation from actual losses to evaluate models’ capabilities in realistically assessing risks.

5.1. SENSITIVITY ANALYSIS

We examine the model’s sensitivity to parameter choices. Specifically, we examine the effect of the value of γ and the demand-damping rate m . Recall that γ is the parameter controlling the conservativeness of the CLT uncertainty set, given by equation (29), and the demand-damping rate controls how fast demand declines in response to increases in price, given by equation (16). To examine the overall perfor-

Table 1. Summary of models implemented for insurance premium calculation.

Model Name	Description
Hist	Uses the historical level of insurance premium collected by the NFIP, relying on past data to set future premiums without adjustment for future uncertainties.
CMA	Employs a CMA to compute insurance premiums using historical losses, aiming for stability through past data analysis.
RO	Implements robust optimization with a central limit theorem, or CLT, uncertainty set, focusing on mitigating risk within a specific range of uncertainty based on statistical theory.
ARO	Adaptive robust optimization that dynamically adjusts insurance premiums in response to changing conditions and uncertainties, optimizing strategies over time to minimize risk.

mance of the premium, we compute the cumulative surplus S across all states over the testing period as follows:

$$S(\gamma) = \sum_{i=1}^N \sum_{t=1}^{T1} p_{i,t} - \sum_{i=1}^N \sum_{t=1}^{T1} l_{i,t}^{hist}, \quad (18)$$

where $l_{i,t}^{hist}$ denotes the historical loss that occurred at location i at time t .

Figure 2 shows the level of surplus as a function of different γ values, with different demand-damping rates. The two dotted lines show the constant surplus computed by two baselines: using the actual premiums collected during this period and the CMA rule. We observe that both baselines incur a loss over the testing period, with historical premiums resulting in about a \$20 billion loss, and the CMA rule resulting in an \$8 billion loss.

We let γ take values between 0 and 1.5 at a step size of 0.1, and we resolve the optimization model at each γ value and compute the sum of surplus across all states across testing years. $\gamma = 0$ corresponds to convex optimization without uncertainty, and $\gamma = 1.5$ corresponds to the maximum degree of uncertainty. As we increase the value of γ , the size of the uncertainty set increases and the model becomes more conservative, resulting in a surplus as expected. We remark that the surplus breaks even when γ takes a value between 0.6 and 0.7. And we observe a smooth increase in surplus as γ increases.

In addition, we experiment with three demand-damping rates: no damping, $m_1 = 1/P_{hist}^{max}$, and $m_2 = 1/(2 * P_{hist}^{max})$, with m_2 damping twice as fast as m_1 . Similar to as above, we experiment with varying γ values corresponding to the different demand-damping rates. We observe that the choice of demand damping is less significant compared with the variation of γ .

5.2. EFFECTIVENESS

Regulators require insurers to protect against insolvency risks, ensuring that premium structures are robust enough to cover potential losses. Therefore, evaluating the effectiveness of insurance schemes in mitigating insolvency is critical. We assess insolvency at the state level, defining a state as insolvent when its cumulative premiums do not cover cumulative losses. This serves as a key indicator of a scheme's financial resilience and risk management efficiency.

Our analysis spans the 2013–2022 testing period, corresponding to the last 10 years of available data, and focuses

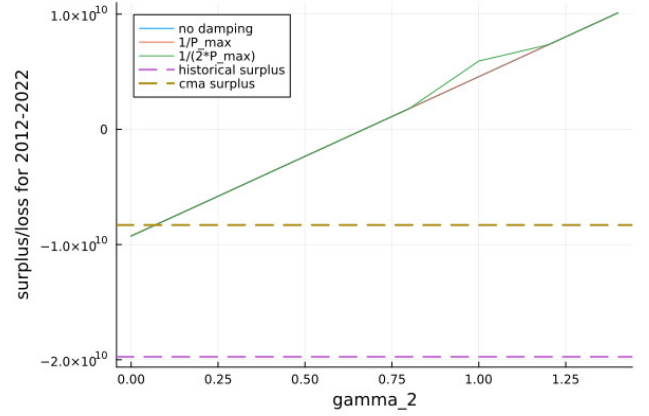


Figure 2. Surplus (or loss) computed for 2012 to 2022 across all states when varying the level of γ . The two dotted lines demonstrate the level of surplus from two baseline models: Historical surplus is calculated from the actual premiums collected, and CMA surplus is computed using the cumulative moving average.

on the impact of γ , a parameter that significantly influences model outcomes as indicated in the sensitivity analysis section. As summarized in Table 2, we examine the effects of varying γ from 0 to 2 in increments of 0.2, evaluating its role in determining the number of insolvent states under each model.

The results in Table 2 highlight substantial differences in the effectiveness of various premium-setting approaches in mitigating insolvency risk. The historical (Hist) scheme performs the worst, resulting in insolvency across all 52 states, underscoring its failure to provide adequate financial protection against catastrophic losses. This outcome reinforces the urgent need for reform in NFIP premium setting in ensuring long-term solvency. The CMA approach, while demonstrating some improvement, still results in insolvency in 36 states. This indicates that although it offers greater stability than the Hist scheme, it lacks the adaptability required to account for evolving risk dynamics. Both the Hist and CMA schemes remain unchanged across different γ values, as expected.

Both the RO and ARO models demonstrate a clear downward trend in insolvency rates as γ increases, reflecting their ability to enhance financial resilience through more conservative parameter choices in the uncertainty set. At γ

Table 2. The number of insolvent states during the testing period based on different methods. In total, we test across 52 states and territories—the 50 US states plus Puerto Rico and the US Virgin Islands.

γ	ARO	RO	CMA	Hist
0.0	36.0	36.0	36.0	52.0
0.2	30.0	32.0	36.0	52.0
0.4	26.0	26.0	36.0	52.0
0.6	25.0	26.0	36.0	52.0
0.8	22.0	23.0	36.0	52.0
1.0	21.0	22.0	36.0	52.0
1.2	19.0	21.0	36.0	52.0
1.4	16.0	18.0	36.0	52.0
1.6	15.0	15.0	36.0	52.0
1.8	13.0	13.0	36.0	52.0
2.0	12.0	12.0	36.0	52.0

= 0, these models perform similarly to the CMA approach, with 36 insolvent states, suggesting that without additional conservatism, they do not provide superior risk protection. However, as γ increases, insolvency rates decline significantly, with ARO consistently outperforming RO across all γ levels. For instance, at $\gamma = 1.0$, ARO reduces the number of insolvent states to 21, compared to 22 for RO, and at $\gamma = 2.0$, ARO achieves the best performance with just 12 insolvent states, matching the lowest observed insolvency rate. The superior performance of ARO over RO suggests that its adaptive component allows for a more refined approach to premium adjustments under uncertainty. By dynamically incorporating new information, ARO ensures a more robust financial buffer against catastrophic events. The results indicate that higher γ values lead to more conservative pricing, directly translating into lower insolvency rates. This trade-off underscores the importance of selecting an appropriate γ value based on risk tolerance and financial strategy. Ultimately, ARO emerges as the most effective method for balancing risk management and capital efficiency, demonstrating its ability to navigate the complexities of catastrophe insurance pricing under uncertainty.

5.3. EFFICIENCY

To evaluate the efficiency of the models, we examine both surplus and deficit levels as well as the absolute deviation (AD) from actual losses. This ensures that models not only mitigate insolvency risks but also avoid excessive overpricing, aligning premiums more accurately with underlying risks.

Table 3 presents the overall surplus (or deficit) levels and AD across all states over the testing period. The AD serves as an indicator of how well each model captures actual risk levels and is computed as a function of γ as follows:

$$AD(\gamma) = \sum_{i=1}^N \sum_{t=1}^{T1} |p_{i,t} - l_{i,t}^{act}|. \quad (19)$$

First, we observe that premium setting under the historical scheme significantly undercharges over the testing period, resulting in \$19 billion in losses, while the CMA rule

results in \$8 million in losses. This suggests that historical levels are insufficient to cover future losses. Similar to as before, we observe that as γ increases, the RO and ARO schemes increase the level of conservatism and achieve more surplus. With the same level of γ , we observe that ARO achieves the surplus more slowly than the RO schemes. Especially when γ reaches the level of 1.4, the ARO scheme increases premiums much more slowly than the RO schemes, which increases at a constant rate with increasing γ . Figure 3 visualizes the efficient frontier, illustrating the trade-off between the number of insolvent states versus the surplus (or deficit) achieved during the testing period.

Both the RO and ARO models exhibit a clear trend of increasing surplus as γ rises, demonstrating their ability to enhance financial resilience by incorporating greater conservatism in pricing. At $\gamma = 0$, RO and ARO perform similarly to CMA, both exhibiting a deficit of approximately \$9.2 billion, indicating that without additional conservatism, these models do not provide better financial protection. However, as γ increases, ARO and RO gradually transition from a deficit to surplus. For instance, at $\gamma = 1.0$, the ARO scheme achieves a surplus of \$0.55 billion, while RO reaches \$4.57 billion. At the highest level, $\gamma = 2.0$, RO achieves a surplus of \$18.41 billion, compared to \$3.82 billion for ARO. RO accumulates surplus at a significantly higher rate than ARO, suggesting that ARO adjusts premiums more gradually, while RO adopts a more aggressive pricing strategy. This difference arises because ARO premiums are influenced by the previous year's losses, thus allowing for greater flexibility and a more measured response to risk fluctuations. A key distinction between ARO and RO emerges as γ increases, particularly at $\gamma = 1.4$, where RO reaches a \$10.11 billion surplus, while ARO remains at \$2.16 billion. This suggests that ARO incorporates uncertainty more dynamically, whereas RO increases premiums at a near-linear rate with γ . This controlled adjustment process in ARO makes it preferable for policymakers seeking a stable and adaptable pricing framework. In contrast, RO provides a more aggressive pricing approach, accumu-

Table 3. Surplus/deficit (S/D) and absolute deviation (AD) across different γ levels for ARO, RO, CMA, and Hist.

γ	ARO		RO		CMA		Hist	
	S/D	AD	S/D	AD	S/D	AD	S/D	AD
0.0	-9.23e9	2.26e10	-9.27e9	2.25e10	-8.31e9	2.30e10	-1.98e10	1.98e10
0.2	-5.88e9	2.47e10	-6.50e9	2.42e10	-8.31e9	2.30e10	-1.98e10	1.98e10
0.4	-3.56e9	2.65e10	-3.73e9	2.59e10	-8.31e9	2.30e10	-1.98e10	1.98e10
0.6	-1.80e9	2.78e10	-0.96e9	2.77e10	-8.31e9	2.30e10	-1.98e10	1.98e10
0.8	-0.46e9	2.88e10	1.80e9	2.95e10	-8.31e9	2.30e10	-1.98e10	1.98e10
1.0	0.55e9	2.96e10	4.57e9	3.13e10	-8.31e9	2.30e10	-1.98e10	1.98e10
1.2	1.41e9	3.03e10	7.34e9	3.32e10	-8.31e9	2.30e10	-1.98e10	1.98e10
1.4	2.16e9	3.09e10	10.11e9	3.53e10	-8.31e9	2.30e10	-1.98e10	1.98e10
1.6	2.78e9	3.13e10	12.88e9	3.74e10	-8.31e9	2.30e10	-1.98e10	1.98e10
1.8	3.32e9	3.18e10	15.64e9	3.96e10	-8.31e9	2.30e10	-1.98e10	1.98e10
2.0	3.82e9	3.21e10	18.41e9	4.17e10	-8.31e9	2.30e10	-1.98e10	1.98e10

lating a higher surplus but at the cost of higher absolute deviation, which may indicate overpricing.

The AD metric provides insights into how well each model aligns premiums with actual incurred losses. The historical scheme has the lowest AD (\$1.98e10), but that is due to severe underpricing rather than accurate risk estimation, as evidenced by its extreme deficit. The CMA scheme maintains an AD of \$2.30e10. Both the RO and ARO schemes lead to an increase in AD as γ increases, while they become more conservative. They also tend to deviate more from actual losses. However, ARO consistently maintains a lower AD than RO, highlighting that ARO balances conservatism with pricing accuracy more effectively.

The trade-off between insolvency risk and surplus generation is further illustrated in [Figure 3](#), which depicts the number of insolvent states against the surplus/deficit levels. The results indicate that while increasing γ improves financial resilience, it also leads to higher absolute deviation from actual losses. ARO emerges as the most balanced scheme, providing strong insolvency protection while avoiding excessive overpricing. RO, on the other hand, is better suited for highly risk-averse insurers who prioritize surplus accumulation over pricing efficiency. The Hist and CMA schemes fail to provide sustainable financial protection, reinforcing the necessity of adopting dynamic, uncertainty-aware premium-setting strategies.

In conclusion, the results confirm that ARO is the most effective model for balancing financial stability, premium efficiency, and risk alignment. Although RO achieves the highest surplus, it does so at the cost of higher deviation, making it more suitable for risk-averse regulators. In contrast, ARO provides a more stable pricing strategy with controlled surplus accumulation, making it a more pragmatic choice for long-term catastrophe insurance pricing. [Figure 7](#), included in the appendix, further explores the trade-off between the number of insolvent states and absolute deviation during the testing period.

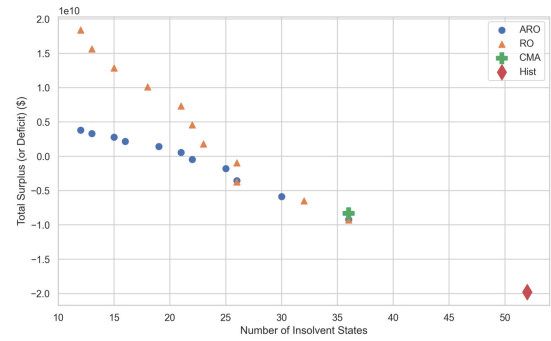


Figure 3. Scatterplot visualizing the efficient frontier, showing how different values of γ affect the number of insolvent states (x-axis) and the total surplus or deficit (y-axis) computed as the total premium charged minus actual loss over the testing period. Note that the CMA and Hist schemes are plotted as static points because their values do not change with varying γ values.

6. CONCLUSION

In this work, we present an ARO framework for catastrophe insurance premium pricing, designed to balance insolvency risk and capital efficiency. We first introduce a nominal linear optimization formulation to frame the challenge of setting insurance prices for rare catastrophic events. We then extend this to an RO model incorporating a central limit theorem uncertainty set, ensuring protection against volatility in losses modeled from historical distributions. By solving the inner problem in closed form, we derive the robust counterpart and reformulate it as a convex optimization problem. Further, we develop an ARO model using linear decision rules allowing for dynamic premium adjustments based on actual loss experiences, improving responsiveness to evolving risk conditions.

We apply our ARO framework to the realm of US flood insurance, leveraging data from the NFIP. We evaluate perfor-

mance against two benchmark schemes—NFIP premiums and CMA premiums. Historical data from 1975 to 2012 were used to construct uncertainty sets, and the last 10 years of NFIP data were used to assess out-of-sample performance. Our results demonstrate that optimization-based approaches effectively cover losses at the same time as they provide flexibility to policymakers in adjusting premiums based on risk tolerance and degree of conservatism.

Our findings point to two key advantages of the ARO framework:

- **Effectiveness:** The ARO approach significantly reduces insolvency risk by effecting a smooth transition from a high number of insolvent states to a low and controllable level as conservatism increases. Unlike static benchmarks, which fail to adapt to risk variations, the ARO scheme provides policy providers with flexibility to determine the desired insolvency level while maintaining financial stability.

- **Efficiency:** The ARO scheme efficiently adjusts premiums to strike a balance between financial sustainability and affordability. Unlike RO, which aggressively increases surplus at increasingly conservative levels, ARO maintains a more controlled surplus accumulation, avoiding excessive overpricing. The model enables a smooth transition from deficit to surplus, ensuring that insurers are neither overexposed to risk nor excessively charging policyholders.

The ARO framework is adaptable and generalizable to a broad class of catastrophic events beyond floods, including wildfire, drought, and extreme weather events. ARO's flexibility makes it a promising tool for insurers and policymakers who wish to develop sustainable, data-driven pricing strategies for managing climate- and disaster-related risks.

Submitted: September 21, 2025 EDT. Accepted: July 08, 2026 EDT. Published: September 14, 2026 EDT.

REFERENCES

- Alem, Douglas, Alistair Clark, and Alfredo Moreno. 2016. "Stochastic Network Models for Logistics Planning in Disaster Relief." *European Journal of Operational Research* 255 (1): 187–206.
- Ben-Tal, Aharon, Byung Do Chung, Supreet Reddy Mandala, and Tao Yao. 2011. "Robust Optimization for Emergency Logistics Planning: Risk Mitigation in Humanitarian Relief Supply Chains." *Transportation Research Part B: Methodological* 45 (8): 1177–89.
- Bertsimas, Dimitris, David B. Brown, and Constantine Caramanis. 2011. *Robust Optimization*. Princeton University Press.
- Bertsimas, Dimitris, and Dick den Hertog. 2022. *Robust and Adaptive Optimization*.
- Charpentier, Arthur, and Benoît Le Maux. 2014. "Natural Catastrophe Insurance: How Should the Government Intervene?" *Journal of Public Economics* 115: 1–17.
- Chen, Weiye, and Limao Zhang. 2022. "An Automated Machine Learning Approach for Earthquake Casualty Rate and Economic Loss Prediction." *Reliability Engineering & System Safety* 225: 108645.
- Dlugolecki, A., C. Crerar, and K. Bermingham. 2009. "Climate Change and Its Implications for Catastrophe Modelling." *Coping with Climate Change: Risks and Opportunities for Insurers*, 1–23.
- "EM-DAT The International Disaster Dataset." 2021. Center for Research on Epidemiology of Disasters. <https://www.emdat.be/>.
- Ermolieva, T., T. Filatova, Y. Ermoliev, M. Obersteiner, K. M. de Bruijn, and A. Jeuken. 2017. "Flood Catastrophe Model for Designing Optimal Flood Insurance Program: Estimating Location-Specific Premiums in the Netherlands." *Risk Analysis* 37 (1): 82–98. <https://doi.org/10.1111/risa.12589>.
- Federal Emergency Management Agency (FEMA). 2023. "OpenFEMA Dataset: FIMA NFIP Redacted Policies - V1."
- FEMA. 2023. "OpenFEMA Dataset: FEMA NFIP Redacted Claims - V2."
- Flavelle, Christopher, Jill Cowan, and Ivan Penn. 2023. "Climate Shocks Are Making Parts of America Uninsurable. It Just Got Worse." *The New York Times*, May. <https://www.nytimes.com/2023/05/31/climate/climate-change-insurance-wildfires-california.html>.
- Grossi, Patricia. 2005. *Catastrophe Modeling: A New Approach to Managing Risk*. Vol. 25. Springer Science & Business Media.
- Horn, Diane P., and Jared T. Brown. 2017. *Introduction to the National Flood Insurance Program (Nfip)*. Congressional Research Service, Washington, DC, USA.
- Jaffee, Dwight M., and Thomas Russell. 1997. "Catastrophe Insurance, Capital Markets, and Uninsurable Risks." *Journal of Risk and Insurance*, 205–30.
- Keucheyan, Razmig. 2018. "Insuring Climate Change: New Risks and the Financialization of Nature." *Development and Change* 49 (2): 484–501.
- Kunreuther, Howard, Geoffrey Heal, Myles Allen, Ottmar Edenhofer, Christopher B. Field, and Gary Yohe. 2013. "Risk Management and Climate Change." *Nature Climate Change* 3 (5): 447–50. <https://doi.org/10.1038/nclimate1740>.
- Linnerooth-Bayer, Joanne, and Stefan Hochrainer-Stigler. 2015. "Financial Instruments for Disaster Risk Management and Climate Change Adaptation." *Climatic Change* 133: 85–100.
- Liu, Tongxin, Jianfang Shao, and Xihui Wang. 2022. "Funding Allocations for Disaster Preparation Considering Catastrophe Insurance." *Socio-Economic Planning Sciences* 84: 101413.
- Mercer, Jessica. 2010. "Disaster Risk Reduction or Climate Change Adaptation: Are We Reinventing the Wheel?" *Journal of International Development: The Journal of the Development Studies Association* 22 (2): 247–64.
- Michel-Kerjan, Erwann, and Howard Kunreuther. 2011. "Redesigning Flood Insurance." *Science* 333 (6041): 408–9. <https://doi.org/10.1126/science.1202616>.
- Michel-Kerjan, Erwann O. 2010. "Catastrophe Economics: The National Flood Insurance Program." *Journal of Economic Perspectives* 24 (4): 165–86.
- Mills, Evan. 2007. "Synergisms between Climate Change Mitigation and Adaptation: An Insurance Perspective." *Mitigation and Adaptation Strategies for Global Change* 12 (5): 809–42. <https://doi.org/10.1007/s11027-007-9101-x>.
- Mitchell-Wallace, Kirsten, Matthew Jones, John Hillier, and Matthew Foote. 2017. *Natural Catastrophe Risk Management and Modelling: A Practitioner's Guide*. John Wiley & Sons.
- Nicholson, Jack E, Karen Clark, and Glen Daraskevich. 2018. "The Florida Insurance Market: An Analysis of Vulnerabilities to Future Hurricane Losses." *Journal of Insurance Regulation* 37 (3): 57–89.
- Olivieri, Annamaria, and Ermanno Pitacco. 2015. *Introduction to Insurance Mathematics: Technical and Financial Features of Risk Transfers*. Springer.
- Rueden, Laura von, Sebastian Mayer, Rafet Sifa, Christian Bauckhage, and Jochen Garcke. 2020. "Combining Machine Learning and Simulation to a Hybrid Modelling Approach: Current and Future Directions." *Advances in Intelligent Data Analysis XVIII: 18th International Symposium on Intelligent Data Analysis, IDA 2020, Konstanz, Germany, April 27–29, 2020, Proceedings* 18, 548–60.
- Salmerón, Javier, and Aruna Apte. 2010. "Stochastic Optimization for Natural Disaster Asset Prepositioning." *Production and Operations Management* 19 (5): 561–74.
- Schneider, Philip J, and Barbara A Schauer. 2006. "HAZUS—Its Development and Its Future." *Natural Hazards Review* 7 (2): 40–44.

- Van Aalst, Maarten K. 2006. "The Impacts of Climate Change on the Risk of Natural Disasters." *Disasters* 30 (1): 5–18.
- Wang, Duo, Kai Yang, and Lixing Yang. 2023. "Risk-Averse Two-Stage Distributionally Robust Optimisation for Logistics Planning in Disaster Relief Management." *International Journal of Production Research* 61 (2): 668–91.
- Yang, Weichao, Kui Xu, Chao Ma, et al. 2021. "A Novel Multi-Objective Optimization Framework to Allocate Support Funds for Flash Flood Reduction Based on Multiple Vulnerability Assessment." *Journal of Hydrology* 603: 127144.
- Zahura, Faria T., Jonathan L. Goodall, Jeffrey M. Sadler, Yawen Shen, Mohamed M. Morsy, and Madhur Behl. 2020. "Training Machine Learning Surrogate Models from a High-Fidelity Physics-Based Model: Application for Real-Time Street-Scale Flood Prediction in an Urban Coastal Community." *Water Resources Research* 56 (10): e2019WR027038.
- Zokaee, Shiva, Ali Bozorgi-Amiri, and Seyed Jafar Sadjadi. 2016. "A Robust Optimization Model for Humanitarian Relief Chain Design under Uncertainty." *Applied Mathematical Modelling* 40 (17–18): 7996–8016.

APPENDICES

A. TECHNICAL PROOFS

A.1. PROOF OF PROPOSITION 3.1

Proof. Since uncertain variables occur only in one constraint, we can write Problem (8) as a min-max problem as follows:

$$\begin{aligned} \min_{p_{i,t}} \max_{l_{i,t} \in \mathcal{U}_i} & \sum_{i=1}^N \sum_{t=1}^T f(p_{i,t}) \cdot p_{i,t} \\ & \sum_{t=1}^T f(p_{i,t}) \cdot p_{i,t} - \sum_{t=1}^T f(p_{i,t}) \cdot l_{i,t} \geq \delta, \quad i \in [N], \\ & |p_{i,t} - p_{i,t-1}| \leq \theta, \quad i \in [N], t \in [T], \\ & (l_{i,t})_{t \in T} \in \mathcal{U}_i, \quad i \in [N], \\ & p_{i,t} \in \mathbb{R}^+, \quad i \in [N], t \in [T]. \end{aligned} \quad (20)$$

Next, we consider the inner problem:

$$\begin{aligned} \max_{l_{i,t} \in \mathcal{U}_i} & \sum_{i=1}^N \sum_{t=1}^T f(p_{i,t}) \cdot p_{i,t} \\ & \sum_{t=1}^T f(p_{i,t}) \cdot p_{i,t} - \sum_{t=1}^T f(p_{i,t}) \cdot l_{i,t} \geq \delta, \quad i \in [N], \\ & (l_{i,t})_{t \in T} \in \mathcal{U}_i, \\ & p_{i,t} \in \mathbb{R}^+, \quad i \in [N], t \in [T]. \end{aligned} \quad (21)$$

For the inner problem, treating $p_{i,t}$ as constants, the objective function is constant. Thus, looking at the constraints containing $l_{i,t}$, the inner problem can be simplified to

$$\begin{aligned} \max_{l_{i,t} \in \mathcal{U}_i} & \sum_{i=1}^N \sum_{t=1}^T l_{i,t} \\ & (l_{i,t})_{t \in T} \in \mathcal{U}_i, \\ & p_{i,t} \in \mathbb{R}^+, \quad i \in [N], t \in [T]. \end{aligned} \quad (22)$$

We observe that the objective involves a summation over $i \in N$, allowing us to decompose the problem into independent subproblems—one for each location. Without loss of generality, we omit the demand-damping curve, since the piecewise linear formulation does not alter the optimal solution. The overall maximum is therefore the sum of the individual maxima across all i . Focusing on a single location, we proceed by substituting the original expression for $\mathcal{U}_i$:

$$\begin{aligned} \max_{l_{i,t}} & \sum_{i=1}^N \sum_{t=1}^T l_{i,t} \\ & \left| \sum_{t=1}^T l_{i,t} - T \cdot \bar{l}_i \right| \leq \gamma \cdot \sigma_i \sqrt{T} \quad \forall i \in [N], \\ & l_{i,t} \in \mathbb{R}^+. \end{aligned} \quad (23)$$

Looking at the constraint, we can take out $|\cdot|$ since we are maximizing over $l_{i,t}$, and rearranging terms gives us

$$\sum_{t=1}^T l_{i,t} \leq T \cdot \bar{l}_i + \gamma \cdot \sigma_i \sqrt{T} \quad \forall i \in [N], \quad (24)$$

Optimality is achieved at equality. We can solve for, i.e., $l_{i,t}^*$ s.t.

$$\sum_{t=1}^T l_{i,t} = T \cdot \bar{l}_i + \gamma \cdot \sigma_i \sqrt{T} \quad \forall i \in [N]. \quad (25)$$

Hence, we can compute for all locations $i \in [N]$ and denote the analytical solution to each subproblem as L_i :

$$L_i := \sum_{t=1}^T l_{i,t}^* = T \cdot \bar{l}_i + \gamma \cdot \sigma_i \sqrt{T} \sqrt{(T_1 - T_0)}. \quad (26)$$

Plugging the analytical solution from the inner problem back to the outer problem, we have the sum of future loss for each location given as follows:

$$\sum_{t=1}^T l_{i,t}^* = T \cdot \bar{l}_i + \gamma \cdot \sigma_i \sqrt{T}. \quad (27)$$

Therefore, the robust counterpart of the original problem is

$$\begin{aligned} \min_{p_{i,t}} & \sum_{i,t} f(p_{i,t}) * p_{i,t} \\ & \sum_{t=1}^T f(p_{i,t}) * p_{i,t} - \frac{1}{T} \sum_{t=1}^T f(p_{i,t}) * L_i \geq \delta, \quad \forall i \in [N], \\ & |p_{i,t} - p_{i,t-1}| \leq \theta, \quad \forall i \in [N], \quad t \in [T], \end{aligned} \quad (28)$$

where

$$L_i = T \cdot \bar{l}_i + \gamma \cdot \sigma_i \sqrt{T}. \quad (29)$$

A.2. PROOF OF PROPOSITION 3.2

Proof. Recall that this optimization problem aims to minimize the total premiums over all locations and time periods while ensuring that the cumulative premium exceeds the cumulative losses by at least a margin of δ . Note that in this formulation, we impose a slowly varying constraint on $\alpha_{i,t}$, instead of $p_{i,t}$ as in the previous formulation, because $p_{i,t}$ now depends on uncertain variables $l_{i,t}$.

Similar to as before, notice that the constraints for each location i are independent of other locations, and thus minimizing the aggregated premium is the same as minimizing the premium for each location. Therefore we can decompose the problem by solving for each location i the ARO problem independently. In addition, noticing that the objective function now depends on the uncertain variable $l_{i,t}$, we therefore use the epigraph formulation as discussed in Chapter 2 of Bertsimas and Hertog (2022) to move all variables containing uncertain variables into constraints. We arrive at the following ARO formulation for one location i :

$$\min_{\alpha_{i,t}, \beta_{i,t}} \Omega \quad (30)$$

$$\sum_{t=1}^T p_{i,t} \leq \Omega, \quad \forall (l_{i,t})_{t \in T} \in \mathcal{U}_i, \quad (31)$$

$$\sum_{t=1}^T p_{i,t} - \sum_{t=1}^T l_{i,t} \geq \delta, \quad \forall (l_{i,t})_{t \in T} \in \mathcal{U}_i, \quad (32)$$

$$\|\alpha_{i,t} - \alpha_{i,t-1}\| \leq \theta_1, \quad t = 2, \dots, T, \quad (33)$$

$$\|\beta_{i,t} - \beta_{i,t-1}\| \leq \theta_2, \quad t = 2, \dots, T, \quad (34)$$

$$p_{i,t} \geq 0, \quad t = 1, \dots, T, \quad \forall (l_{i,t})_{t \in T} \in \mathcal{U}_i, \quad (35)$$

where $p_{i,t}$ is a quantity that depends on the uncertain losses given by equation (10). Next, we derive the robust counterpart by taking the robust counterpart for each constraint that depends on the uncertain variable $l_{i,t}$ independently.

First, consider the constraint given by inequality (31). Replacing $p_{i,t}$ with $\alpha_{i,t}, \beta_{i,t}$, which now depends on the uncertain variable $l_{i,t}$, we rewrite the constraint as follows

$$\sum_{t=1}^T \alpha_{i,t} + \sum_{t=2}^T \beta_{i,t} \cdot l_{i,t-1} \leq \Omega, \quad \forall (l_{i,t})_{t \in T} \in \mathcal{U}_i, \quad (36)$$

which is is equivalent to

$$\sum_{t=1}^T \alpha_{i,t} + \max_{l_{i,t} \in \mathcal{U}_i} \left\{ \sum_{t=2}^T \beta_{i,t} \cdot l_{i,t-1} \right\} \leq \Omega. \quad (37)$$

Let's now consider the inner maximization problem, using the explicit expression for $\mathcal{U}_i^{CLT}$. We have a linear programming (LP) problem in $l_{i,t}$:

$$\begin{aligned} & \max_{l_{i,t}} \sum_{t=2}^T \beta_{i,t} \cdot l_{i,t-1} \\ & \sum_{t=1}^T l_{i,t} \leq \bar{T}_i + \gamma \cdot \sigma_i \cdot \sqrt{T}, \\ & -\sum_{t=1}^T l_{i,t} \leq -\bar{T}_i + \gamma \cdot \sigma_i \cdot \sqrt{T}. \end{aligned} \quad (38)$$

The inner problem thus satisfies strong duality. Notice that the uncertainty set is given in polyhedral form. We can thus take the dual by introducing the dual variables s_1^1, s_2^1 , and we arrive at the following form:

$$\begin{aligned} & \min_{s_1^1, s_2^1} \sum_{j=1}^2 c_j s_j^1 \\ & s_1^1 - s_2^1 \geq \beta_t, \quad \forall t = 2, \dots, T, \\ & s_1^1 - s_2^1 \geq 0, \\ & s_1^1, s_2^1 \geq 0. \end{aligned} \quad (39)$$

By strong duality, the inner maximization problem has the same objective value of the dual minimization problem. Using the dual expression, constraint (31) becomes

$$\begin{aligned} & \sum_{t=1}^T \alpha_{i,t} + \min_{s_1^1, s_2^1} \sum_{j=1}^2 c_j s_j^1 \leq \Omega \\ & s_1^1 - s_2^1 \geq \beta_{i,t}, \quad \forall t = 2, \dots, T, \\ & s_1^1 - s_2^1 \geq 0, \\ & s_1^1, s_2^1 \geq 0. \end{aligned}$$

Note that we can take away the minimization term because if any feasible s_1^1, s_2^1 satisfies this constraint, the minimum also does.

Second, consider the constraint given by inequality (32), which ensures premiums cover losses:

$$\sum_{t=1}^T p_{i,t} - \sum_{t=1}^T l_{i,t} \geq \delta, \quad \forall (l_{i,t})_{t \in T} \in \mathcal{U}_i. \quad (40)$$

Multiplying both sides by -1 , we have

$$\sum_{t=1}^T l_{i,t} - \sum_{t=1}^T p_{i,t} \leq -\delta. \quad (41)$$

Substituting $\alpha_{i,t}$ and $\beta_{i,t}$ for $l_{i,t}$,

$$\sum_{t=1}^T l_{i,t} - \left[\sum_{t=1}^T \alpha_{i,t} + \sum_{t=2}^T \beta_{i,t} l_{i,t-1} \right] \leq -\delta, \quad (42)$$

and rearranging terms to collect all the uncertain terms $l_{i,t}$,

$$-\sum_{t=1}^T \alpha_{i,t} + \left[l_{i,1} + \sum_{t=2}^T (1 - \beta_{i,t}) l_{i,t} \right] \leq -\delta, \quad (43)$$

we finally arrive at

$$-\sum_{t=1}^T \alpha_{i,t} + \max_{l_{i,t} \in \mathcal{U}_i} \left\{ l_{i,1} + \sum_{t=2}^T (1 - \beta_{i,t}) l_{i,t} \right\} \leq -\delta. \quad (44)$$

We now consider the inner maximization problem, which is once again an LP in $l_{i,t}$:

$$\begin{aligned} & \max_{l_{i,t}} l_{i,1} + \sum_{t=2}^T (1 - \beta_{i,t}) \cdot l_{i,t} \\ & \sum_{t=1}^T l_{i,t} \leq \bar{T}_i + \gamma \cdot \sigma_i \cdot \sqrt{T}, \\ & -\sum_{t=1}^T l_{i,t} \leq -\bar{T}_i + \gamma \cdot \sigma_i \cdot \sqrt{T}. \end{aligned} \quad (45)$$

Similar to as before, by strong duality, the dual is given as a minimization problem in dual variables s_1^2, s_2^2 :

$$\begin{aligned} & \min_{s_1^2, s_2^2} \sum_{j=1}^2 c_j s_j^2 \\ & s_1^2 - s_2^2 \geq 1 - \beta_{i,t}, \quad \forall t = 2, \dots, T, \\ & s_1^2 - s_2^2 \geq 1, \\ & s_1^2, s_2^2 \geq 0. \end{aligned} \quad (46)$$

Therefore constraint (32) becomes

$$\begin{aligned} & -\sum_{t=1}^T \alpha_{i,t} + \sum_{j=1}^2 c_j s_j^2 \leq -\delta, \\ & s_1^2 - s_2^2 \geq 1 - \beta_{i,t}, \quad \forall t = 2, \dots, T, \\ & s_1^2 - s_2^2 \geq 1, \\ & s_1^2, s_2^2 \geq 0. \end{aligned} \quad (47)$$

Finally, we consider the positivity constraint as given by inequality (35):

$$p_{i,t} \geq 0, \quad \forall t \in [T], \quad \forall (l_{i,t})_{t \in T} \in \mathcal{U}_i. \quad (48)$$

Substituting $\alpha_{i,t}, \beta_{i,t}$ for $p_{i,t}$, the constraint is equivalent to

$$\alpha_{i,t} + \beta_{i,t} \cdot l_{i,t-1} \geq 0, \quad \forall t \in [T], \quad \forall (l_{i,t})_{t \in T} \in \mathcal{U}_i. \quad (49)$$

For each time t , we have one separate constraint. Therefore, considering this constraint for some t ,

$$\alpha_{i,t} + \beta_{i,t} \cdot l_{i,t-1} \geq 0, \quad \forall (l_{i,t})_{t \in T} \in \mathcal{U}_i.$$

Multiplying both sides by -1 and rearranging terms, we have

$$-\alpha_{i,t} + \max_{l_{i,t} \in \mathcal{U}} \{-\beta_{i,t} \cdot l_{i,t-1}\} \leq 0.$$

Similar to as before, we consider the inner maximization problem

$$\max_{l_{i,t} \in \mathcal{U}} -\beta_{i,t} \cdot l_{i,t-1}, \quad (50)$$

which has strong duality, with dual given for each t in dual variables $s_1^{3,t}, s_2^{3,t}$ as

$$\begin{aligned} & \min_{s_1^{3,t}, s_2^{3,t}} \sum_j c_j s_j^{3,t} \\ & s_1^{3,t} - s_2^{3,t} \geq -\beta_{i,t} \\ & s_1^{3,t} - s_2^{3,t} \geq 0 \\ & s_1^{3,t}, s_2^{3,t} \geq 0 \end{aligned} \quad (51)$$

Thus, the positivity constraint for some t becomes

$$\begin{aligned} & \alpha_{i,t} + \sum_{j=1}^2 c_j s_j^{3,t} \leq 0 \\ & s_1^{3,t} - s_2^{3,t} \geq -\beta_{i,t} \\ & s_1^{3,t} - s_2^{3,t} \geq 0 \\ & s_1^{3,t}, s_2^{3,t} \geq 0 \end{aligned} \quad (52)$$

Therefore, the overall robust counterpart of the ARO formulation is given as follows:

$$\min_{\alpha_{i,t}, \beta_{i,t}} \Omega \quad (53)$$

with epigraph constraint

$$\begin{aligned}
& \sum_{t=1}^T \alpha_{i,t} + \min_{s_1^1, s_2^1} \sum_{j=1}^2 c_j s_j \leq \Omega, \\
& s_1^1 - s_2^1 \geq \beta_{i,t}, \quad \forall t = 2, \dots, T, \\
& s_1^1 - s_2^1 \geq 0, \\
& s_1^1, s_2^1 \geq 0,
\end{aligned} \tag{54}$$

with loss coverage constraint

$$\begin{aligned}
& - \sum_{t=1}^T \alpha_{i,t} + \sum_j c_j s_j^2 \leq -\delta, \\
& s_1^2 - s_2^2 \geq 1 - \beta_{i,t}, \quad \forall t = 2, \dots, T, \\
& s_1^2 - s_2^2 \geq 1, \\
& s_1^2, s_2^2 \geq 0,
\end{aligned} \tag{55}$$

with positivity constraint

$$\begin{aligned}
& \alpha_{i,t} + \sum_j c_j s_j^{3,t} \leq 0, \\
& s_1^{3,t} - s_2^{3,t} \geq -\beta_{i,t}, \\
& s_1^{3,t} - s_2^{3,t} \geq 0, \\
& s_1^{3,t}, s_2^{3,t} \geq 0,
\end{aligned} \tag{56}$$

and with slowly varying constraint

$$\begin{aligned}
& \|\alpha_{i,t} - \alpha_{i,t-1}\| \leq \theta_1, \quad t = 2, \dots, T, \\
& \|\beta_{i,t} - \beta_{i,t-1}\| \leq \theta_2, \quad t = 2, \dots, T.
\end{aligned} \tag{57}$$

B. OPTIMIZATION MODEL PARAMETER ESTIMATION

In this section, we expand on how to estimate parameters for the optimization model.

B.1. COMPUTING L_i

Recall from equation (29), to compute L_i we need to compute the historical mean and variance for each state. Table 4 shows the historical mean and standard deviation for the top-10 most costly states, computed on an annual basis.

Table 4. Statistics of the top-10 most costly states in the US, including maximum annual claim loss, mean, standard deviation, and median.

State	Max	Mean	Std	Median
LA	3,763,390	45,084	58,467	20,583
TX	8,973,270	44,046	65,420	19,801
NJ	4,022,518	32,727	58,000	13,042
NY	9,467,720	35,725	74,510	13,306
FL	9,100,033	25,351	63,809	8,052
MS	10,000,000	46,603	94,924	14,823
NC	1,294,678	21,725	41,345	7,992
PA	1,889,793	19,040	41,943	6,939
AL	4,900,000	28,059	95,474	8,466
SC	1,764,000	25,574	43,752	10,088

B.2. DEMAND DAMPING

In this study, we model demand sensitivity to insurance premiums through a piecewise linear demand function. We estimate the decline rate using historical data from several states. Figures 4 and 5 show scatterplots of number of policyholders in a year against the mean policy premium of that state at that year. Different states have different degrees of sensitivity to price, but in general we observe a downward trend of declining policyholder numbers as a function of increased price. For the illustrative purpose of this work, we do not specify different sensitivity in different states but use the same demand-damping function across all states.

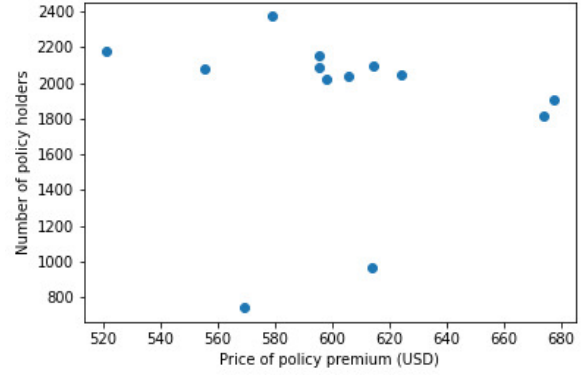


Figure 4. Demand-damping estimation for the state of Louisiana (LA).

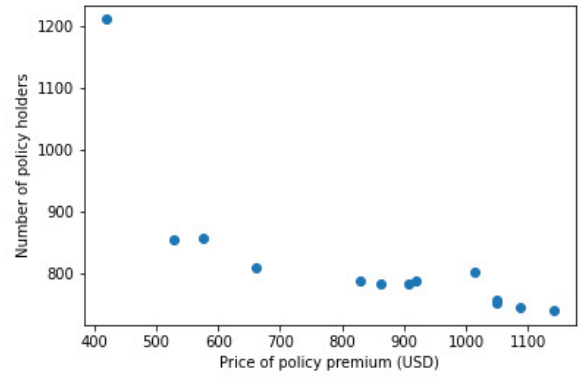


Figure 5. Demand-damping estimation for the state of New York (NY).

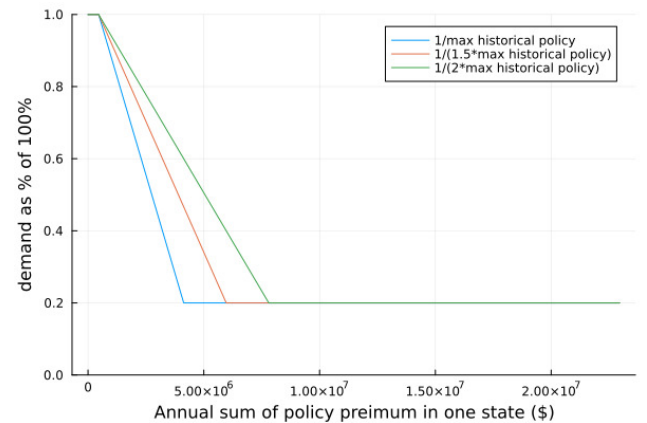


Figure 6. Different piecewise linear demand-damping curves corresponding to different rates of decline.

C. ADDITIONAL PLOT OF THE EFFICIENT FRONTIER

Figure 7 illustrates the trade-off between the number of insolvent states versus the absolute deviation during the testing period.

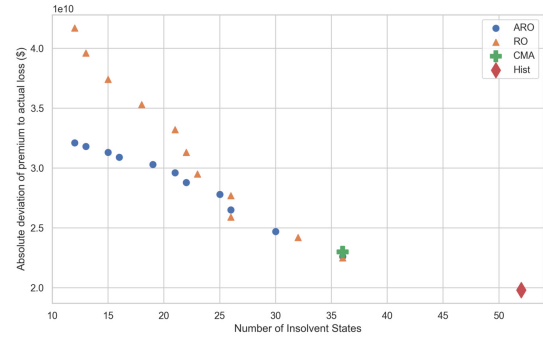


Figure 7. Scatterplot visualizing the efficient frontier, showing how different values of γ affect the number of insolvent states (x-axis) and the total surplus (or deficit) (y-axis) computed as the total premium charged minus actual loss over the testing period. Note that CMA and Hist are plotted as static points because their values do not change with varying γ values. The plot demonstrates that ARO achieves higher efficiency with lower absolute deviation at high γ values, highlighting its better performance under these conditions.