International Journal of Engineering, Science and Environment
Peer-Reviewed Open Access • CC BY 4.0 • Quarterly
Vol. 1 | Issue 1 | Apr–Jun 2026

BOUNDED DATA-DRIVEN ACTUARIAL MAPS AND CVAR-BASED REINSURANCE SELECTION

Nikolaos Halidias

Department of Statistics and Actuarial – Financial Mathematics, University of the Aegean, Samos, Greece

Corresponding Author: nick@aegean.gr

ABSTRACT

Classical actuarial modeling traditionally relies on pre-specified parametric families, an approach that often forces complex data into rigid stochastic structures and introduces significant model risk. In this paper, we advocate for a paradigm shift towards a model-agnostic, data-driven framework. Instead of "marrying" a specific theoretical distribution, we leverage the expressive power of Machine Learning (ML) to fit complex, non-linear patterns directly from data. To ensure actuarial validity, these flexible ML architectures are systematically wrapped in strict admissibility constraints—such as bounded probabilities, positive hazards, and monotone survival curves—ensuring outputs are mathematically rigorous by construction. Furthermore, we challenge the traditional reliance on theoretical "fair" premiums for reinsurance optimization. Recognizing that actual market prices incorporate frictions, capital costs, and risk loadings absent in theoretical models, we formulate reinsurance selection as a practical optimization over a finite-menu of quoted market contracts. By minimizing the sum of the quoted reinsurance premium and the empirical Conditional Value-at-Risk (CVaR) of the retained loss, we bridge the gap between data-driven actuarial modeling and real-world, market-aware tail-risk transfer. This framework provides a robust, implementable methodology for modern risk assessment and capital allocation.

Keywords: Actuarial modeling; sigmoidal approximation; data-driven reserving; survival modeling; ruin probability; CVaR; reinsurance optimization; stop-loss treaties; loss layers.

Mathematics Subject Classification (2020). Primary: 91G05; Secondary: 91G70, 62G05, 62P05, 68T07, 91G60.

1. INTRODUCTION

Actuarial mathematics is fundamentally concerned with the pricing, valuation, reserving, and solvency control of contracts exposed to demographic, underwriting, and financial uncertainty. The classical workflow typically begins by specifying a fully parametric stochastic model for the underlying risk drivers: mortality intensities for lifetimes, frequency–severity models for claims, or stochastic discount factors for valuation. While these models provide analytical tractability and interpretability, they inherently require the actuary to "marry" a specific functional form. This introduces severe model risk: the selected family may be misspecified, too rigid to capture complex non-linear covariate interactions, or poorly calibrated for heterogeneous portfolios. When reality deviates from the chosen parametric assumption, the resulting actuarial valuations can be systematically biased.

The advent of Machine Learning (ML) offers a compelling alternative to this rigid paradigm. ML algorithms enable data-driven estimation of complex functional relationships without requiring pre-specified parametric forms., capturing complex, non-linear patterns without requiring the analyst to pre-specify the underlying stochastic mechanics. However, blindly applying standard, unconstrained ML architectures (such as deep neural networks) to actuarial problems often yields outputs that violate fundamental actuarial axioms. An unconstrained network might predict a survival probability greater than one, a non-monotone ruin probability, or a negative reserve. This paper proposes a conceptual and practical resolution to this dilemma. Our first core contribution is a framework that separates the data-fitting capability of ML from the structural constraints of actuarial science. We demonstrate how flexible, data-driven sigmoidal architectures can be seamlessly wrapped in admissibility constraints. By utilizing specific link functions, convex mixture weights, and bounded parameterizations, we ensure that the ML model retains its universal approximation power while its outputs are guaranteed to be probability-valued, positive, or monotone by construction. The goal is not to invent new mathematical theorems for ML, but to provide a robust, model-agnostic wrapper that allows actuaries to safely deploy modern data science tools without violating the foundational rules of the profession. While individual components draw on established results, the contribution of this paper lies in their systematic integration into a unified actuarial framework.

Our second core contribution addresses a critical blind spot in traditional reinsurance optimization: the persistent disconnect between theoretical pricing and market reality. The classical literature often assumes that reinsurance can be optimized purely based on the expected loss or a theoretical risk measure derived from a fitted probability model. In practice, however, reinsurance premiums are quoted by the market. These market quotes incorporate capital costs, underwriting cycles, counterparty risk, and profit loadings that no internal theoretical model can perfectly capture. Relying solely on theoretical "fair" premiums to select a retention level can lead to severe mispricing and suboptimal risk transfer, as the theoretical cost may often differ from the actual market cost.

To resolve this, we shift the reinsurance selection problem from a theoretical continuous optimization to a practical, market-aware decision process. Given a set of simulated or historical loss scenarios, we evaluate a finite-menu of actual quoted stop-loss contracts. The optimal retention is selected by minimizing the total economic cost: the actual quoted reinsurance premium plus the empirical Conditional Value-at-Risk (CVaR) of the retained loss. This approach directly aligns the actuarial model with the reality of the Chief Financial Officer (CFO) or Chief Actuary, optimizing tail-risk transfer against real-world market quotes rather than theoretical constructs.

The remainder of this paper is structured as follows. Section 2 details the design principles for admissible, data-driven actuarial maps, emphasizing the integration of ML with actuarial constraints. Section 3 outlines core applications in mortality, reserving, and ruin probabilities. Section 4 discusses the theoretical guarantees of the framework and practical ML implementation. Section 5 formalizes the reinsurance market analogy. Section 6 presents the CVaR-based finite-menu selection methodology, accompanied by numerical studies that demonstrate the impact of market quotes versus theoretical pricing. Finally, Section 7 concludes.

2. ADMISSIBLE SIGMOIDAL ACTUARIAL MAPS

Let $S \in K \subset \mathbb{R}^d$ denote observable actuarial features, where K is compact. Depending on the problem, S may include age, duration, calendar year, exposure measures, run-off triangle features, policy characteristics, surplus, attachment points, or macro-financial covariates. Let

$$\sigma(x) = 1 / (1 + e^{-x}),\ \ x \in \mathbb{R},$$

although the discussion applies to any continuous non-decreasing sigmoid with limits 0 and 1 at $-\infty$ and $+\infty$.

A basic shallow sigmoidal network has the form

$$g_\theta(s) = \beta_0 + \Sigma^{m}_{j=1}\, c_j\sigma(w^T_j s + b_j),\ \ s \in K,\ \ (2.1)$$

where $\theta = (\beta_0, c_j, w_j, b_j)^m_{j=1}$. The map g_θ is a flexible real-valued approximation. To obtain actuarially admissible outputs, one should then apply the appropriate link or constraints.

2.1 Probability-valued maps

For lapse, death, default, ruin, or exceedance probabilities, two common constructions are:

$$p_\theta(s) = \sigma(g_\theta(s)),\ \ (2.2)$$

$$p_\theta(s) = \Sigma^{m}_{j=1}\, a_j\sigma(w^T_j s + b_j),\ \ a_j \ge 0,\ \ \Sigma^{m}_{j=1}\, a_j \le 1.\ \ (2.3)$$

The first uses an unconstrained real-valued score followed by a probability link. The second uses a convex-subprobability mixture of sigmoids.

Remark 2.1 *(Probability Admissibility).* The parameterization (2.2) satisfies $0 < p_\theta(s) < 1$ for every $s \in K$ by the range of the sigmoid function. The parameterization (2.3) satisfies $0 \le p_\theta(s) \le 1$ because $0 < \sigma(\cdot) < 1$ and the weights a_j form a convex sub-probability measure. These constructions ensure that the ML model's outputs are mathematically rigorous probabilities by construction, avoiding the common pitfall of unconstrained networks predicting values outside [0, 1].

2.2 Positive hazards and monotone survival functions

For mortality and longevity, a direct probability fit may violate the monotonicity of a survival curve. A safer construction is to model the hazard rate as positive and then integrate it. For an attained age x and future duration $t \ge 0$, define

$$\mu_\theta(x + t, z) = \exp\{g_\theta(x + t, z)\},\ \ (2.4)$$

where z denotes covariates. Then set

$$S_\theta(t \mid x, z) = \exp(-\textstyle\int^t_0 \mu_\theta(x + u, z)\, du).\ \ (2.5)$$

Remark 2.2 *(Admissible Survival Curve Construction).* By defining the hazard rate as strictly positive via the exponential link (2.4), the resulting survival curve (2.5) is guaranteed to satisfy $S_\theta(0 \mid x, z) = 1$, remain in (0, 1], and be non-increasing. This hazard-based wrapper ensures that the flexible ML core g_θ cannot produce actuarially invalid survival curves, regardless of the complexity of the covariate interactions it captures.

2.3 Bounded real-valued maps

For reserves, continuation values, capital add-ons, or discount adjustments, the output may not be a probability but may need to lie in a known interval. If $a < b$, define

$$R_\theta(s) = a + (b - a)\sigma(g_\theta(s)).\ \ (2.6)$$

Then $R_\theta(s) \in (a, b)$ for every $s \in K$. This is useful for technical provisions subject to governance ranges, bounded discount-factor adjustments, bounded loss-development factors, or capped capital allocation rules.

International Journal of Engineering, Science and Environment
Peer-Reviewed Open Access • CC BY 4.0 • Quarterly
Vol. 1 | Issue 1 | Apr–Jun 2026

3. CORE ACTUARIAL APPLICATIONS

3.1 Mortality and survival

Let T_x be the future lifetime of an individual aged x. Under (2.4)–(2.5), the survival probability is estimated as

$$P(T_x > t \mid z) \approx S_\theta(t \mid x, z).$$

The construction incorporates covariates while preserving the essential shape constraints of survival modeling. Death probabilities over one-year intervals can then be derived by

$$q_\theta(x, z) = 1 - S_\theta(1 \mid x, z).$$

3.2 Pricing and valuation

Let X_{tk} denote contract cash flows and $D(t_k)$ discount factors. A classical present value is

$$V(s) = E\left[\sum_k (3.1)\, D(t_k) X_{\{t_k\}} \mid S = s\right] \quad (3.1)$$

A data-driven approximation fits V(s) directly, for example by (2.6) when economic or regulatory bounds are known, or by an unconstrained gθ when the target is real-valued. This can be useful for annuities with guarantees, policyholder behaviour models, and nested simulation replacement.

3.3 Reserving

Let Z denote reserving features extracted from run-off triangles, policy-level claims data, exposure information, and calendar-period covariates. A reserve map can be written as

$$R(Z) = E[\text{outstanding losses} \mid Z]$$

The approximation $R_\theta(Z)=a+(b-a)\sigma(g_\theta(Z))$ enforces positivity and an upper governance bound. Shape restrictions can be added through monotone networks or by constraining selected weights when actuarial expertise implies monotonicity, for example in exposure or development maturity.

3.4 Ruin probabilities and tail-risk maps

Let u denote initial surplus and h a vector of contract or portfolio features. A ruin probability surface may be represented by

$$\psi_\theta(u,t,h)=\sigma(g_\theta(u,t,h)).$$

The final sigmoid ensures that the estimate is a probability. Monotonicity in u should be imposed when required; for example, a higher initial surplus should not increase the ruin probability. Similarly, a portfolio tail-risk map may be fitted as

$$\rho_\theta(h) \approx \rho(L \mid h),$$

where ρ may be VaRα, CVaRα, or a distortion risk measure. When the objective is tail-sensitive, scenario-based optimization is often preferable to purely pointwise regression.

4. GUARANTEES AND IMPLEMENTATION

4.1 Boundedness and Lipschitz stability

Remark 4.1 *(Boundedness and Lipschitz Stability).* Because the sigmoid function is bounded in (0, 1) and has a bounded derivative, it provides a stable output transformation. Furthermore, if the activation function σ is Lipschitz continuous and the network parameters are constrained to a compact set, then the shallow network g_θ is Lipschitz continuous with a bounded Lipschitz constant. Specifically, bounded network parameters ensure that g_θ remains uniformly bounded, while the Lipschitz constant of g_θ is controlled by the norms of the network weights. Consequently, any linked output map $a+(b-a)\sigma(g_\theta(s))$ remains strictly confined to the interval (a, b). For a one-hidden-layer network, the Lipschitz constant is bounded by $\max|c_j| \cdot \|w_j\| \cdot \|\sigma'\|_\infty$. This property ensures that small perturbations in input features do not produce unstable variations in actuarial outputs. in the actuarial output.

4.2 Approximation power

Theorem 4.2 *(Universal approximation; Cybenko 1989; Hornik et al. 1989).*

Let $K \subset \mathbb{R}^d$ be compact and let $h : K \to \mathbb{R}$ be continuous. If σ is a non-polynomial sigmoid, then for every ε > 0 there exist m and parameters θ such that

$$\sup_{s\in K} |g_\theta(s) - h(s)| < \varepsilon.$$

Remark 4.3. The theorem justifies flexible approximation on bounded actuarial domains. It does not, by itself, enforce actuarial admissibility. Probability, positivity, monotonicity, and boundedness must be imposed through the link functions or constraints described above.

4.3 Existence and consistency of empirical fitting

Assumption 4.4 *(Compact parameter set).* For a fixed-architecture size m, the admissible parameter set Θ is compact. This can be enforced by bounding β_0, c_j, w_j, and b_j.

For a given sample, define

$$J_n(\theta) = (1/n) \sum_{i=1}^{n} \ell(f_\theta(S_i), Y_i),$$

and, whenever the argmin is non-empty, denote an empirical minimizer by

$$\hat{\theta}_n \in \arg\min_{\theta\in\Theta} J_n(\theta).$$

The notation $\hat{\theta}_n$ is used below for any such empirical minimizer.

Proposition 4.5 *(Existence and Consistency).* Let $(S_i, Y_i)_{i=1}^{n}$ be data, where f_θ is one of the admissible maps above and ℓ is continuous. Under Assumption 4.4, if $\theta \mapsto J_n(\theta)$ is continuous, then there exists at least one empirical minimizer $\hat{\theta}_n$. Furthermore, under standard data and loss regularity conditions (e.g., i.i.d. observations and an integrable envelope), the empirical risk converges uniformly to the expected risk, ensuring that $J(\hat{\theta}_n) \to \inf_{\theta\in\Theta} J(\theta)$ in probability.

Remark 4.6 *(Limitation and Fixed-architecture).* The preceding result is intentionally fixed-architecture. If the number of hidden units grows with the sample size, or if unbounded losses are used without truncation or envelope control, additional empirical-process assumptions (e.g., Rademacher complexity bounds) are needed. Thus the framework is distribution-light, but not assumption-free.

International Journal of Engineering, Science and Environment
Peer-Reviewed Open Access • CC BY 4.0 • Quarterly
Vol. 1 | Issue 1 | Apr–Jun 2026

4.4 Practical Implementation: Architecture Selection and Regularization

While the theoretical guarantees above assume a fixed-architecture size m, practical implementation requires guidance on model selection and regularization. In practice, the actuary does not need to develop custom optimization routines; the internal parameters (w_j, b_j) can be trained using standard ML frameworks (e.g., PyTorch, TensorFlow) with standard loss functions (e.g., cross-entropy for probabilities, MSE for bounded reserves).

To select the number of hidden units m and prevent overfitting, standard K-fold cross-validation or information criteria (AIC/BIC) adapted for bounded losses can be employed. Importantly, standard regularization techniques (such as L1 or L2 penalties on the inner weights w_j and biases b_j) are fully compatible with the admissibility constraints. Because the actuarial constraints are enforced at the output layer (via the final link function or convex mixture weights), penalizing the inner feature mappings does not violate the boundedness, positivity, or monotonicity of the final output. This separation of concerns—ML handles the data fitting and regularization, while the output layer enforces actuarial axioms—is what makes the framework both flexible and robust. Despite these advantages, it is important to acknowledge that the ML component introduces its own form of model risk. The choice of architecture size m, activation function, and regularization strength can materially affect out-of-sample performance, and outputs may degrade under distribution shift when future loss experience differs from the training data. Actuaries deploying this framework in a regulatory context should therefore subject the fitted model to periodic back-testing, holdout validation, and sensitivity analysis across architecture choices, in line with emerging supervisory guidance on model risk management for insurance applications.

5. REINSURANCE AS TAIL-LAYER TRANSFER

5.1 Treaty notation and option analogy

Let $L \geq 0$ denote an aggregate loss over a fixed horizon. A reinsurance treaty is represented by an indemnity function $I : \mathbb{R}_+ \to \mathbb{R}_+$, where I(L) is paid by the reinsurer to the insurer. The insurer's retained loss is

$$L^I = L - I(L). \quad (5.1)$$

The insurer pays a reinsurance premium $\pi(I)$ to the reinsurer.

Definition 5.1 *(Stop-loss / excess-of-loss treaty).* Fix a retention level $K \geq 0$. The stop-loss (excess-of-loss) indemnity is

$$I^K(L) := (L - K)_+ = \max\{L - K, 0\}.$$

The insurer's retained loss is $L_{net}^K = L - (L - K)_+ = L \wedge K$.

Remark 5.2 *(Loss-Layer Decomposition and Option Analogy).* For any retention $K \geq 0$, the aggregate loss decomposes as $L = (L \wedge K) + (L - K)_+$. The stop-loss indemnity $I^K(L) = (L - K)_+$ is exactly the payoff of a European call option written on the loss variable L with strike K. The insurer is long this call layer (receiving the indemnity), while the reinsurer is short the layer in exchange for the premium. From the insurer's balance-sheet perspective, this call-type payoff caps the retained loss at $L \wedge K$, acting as put-like downside protection for the underwriting result. This option analogy is foundational in financial reinsurance and alternative risk transfer [3, 5, 7].

5.2 Risk-measure and ruin effects

Remark 5.3 *(Bounded Retained Loss and Tail-Risk Reduction).* Under stop-loss reinsurance with retention K, the retained loss $L^K = L \wedge K$ is bounded by K almost surely. Consequently, $L^K \leq L$ pointwise, meaning that every monotone risk functional ρ (including VaRα and CVaRα) satisfies $\rho(L^K) \leq \rho(L)$. Thus, before accounting for the reinsurance premium, the treaty reduces tail risk under monotone risk measures.

Definition 5.4 *(VaR and CVaR).* For a loss X and confidence level $\alpha \in (0, 1)$, let $VaR\alpha(X)$ be an α-quantile of X. Define

$$CVaR\alpha(X) = \inf_{m \in \mathbb{R}} \{ m + (1/(1-\alpha)) E[(X - m)_+] \}.$$

Proposition 5.5 *(One-period ruin comparison with reinsurance premium).* Consider a one-period surplus model with initial surplus $u \geq 0$, premium income $c \geq 0$, aggregate loss L, and stop-loss retention K with reinsurance premium $\pi(K)$. Without reinsurance, $U^0_1 = u + c - L$, whereas with stop-loss reinsurance, $U^K_1 = u + c - \pi(K) - (L \wedge K)$. Hence, ruin occurs if $L > u + c$ without reinsurance, and if $L > u + c - \pi(K)$ with reinsurance (provided $0 \leq u + c - \pi(K) < K$).

Remark 5.6 *(Premium caveat).* The comparison in Proposition 5.5 shows why ruin comparisons cannot be inferred from the pointwise inequality $L \wedge K \leq L$ alone: if the post-premium surplus covers the retention, ruin is eliminated in the one-period model; otherwise the reinsurance premium lowers the ruin threshold. Multi-period ruin comparisons require corresponding assumptions on premium income, renewal dynamics, investment income, treaty renewal, and the net profit condition.

6. CVaR-BASED REINSURANCE SELECTION

A critical limitation of traditional reinsurance optimization is its reliance on theoretical "fair" premiums derived from internal probability models. In practice, reinsurance premiums are quoted by the market and incorporate capital costs,

International Journal of Engineering, Science and Environment
Peer-Reviewed Open Access • CC BY 4.0 • Quarterly
Vol. 1 | Issue 1 | Apr–Jun 2026

underwriting cycles, counterparty risk, and profit loadings that no internal model can perfectly capture. Relying solely on theoretical pricing can lead to severe mispricing, as the theoretical cost will almost always be lower than the actual market cost. Therefore, we shift the reinsurance selection problem to a practical, market-aware decision process: given a set of loss scenarios, we evaluate a finite-menu of actual quoted stop-loss contracts and select the retention that minimizes the total economic cost (quoted premium plus empirical CVaR of the retained loss).

6.1 Scenario data and market quotes

Let $L_1, \ldots, L_n \in \mathbb{R}_+$ be scenario losses. These may be historical annual losses, simulated internal-model outcomes, catastrophe model scenarios, or stress-test outputs. Suppose the reinsurer quotes a finite-menu of stop-loss contracts

$$\mathcal{M} = \{(K_j, \pi_j) : j = 1, \ldots, M\},$$

where K_j is the retention and π_j is the quoted reinsurance premium. For offer j, define retained scenario losses

$$X_{i,j} = L_i \wedge K_j, \quad i = 1, \ldots, n. \quad (6.1)$$

Definition 6.1 *(Empirical CVaR).* For scenario losses $Y_1, \ldots, Y_n$ and $\alpha \in (0, 1)$, define

$$\widehat{CVaR}_\alpha(Y) = \min_{m\in\mathbb{R}} \{ m + (1/((1-\alpha)n)) \Sigma^n_{i=1} (Y_i - m)_+ \}.$$

The scalar m is an empirical VaR proxy. The representation is convex in m and is particularly convenient for optimization.

6.2 Premium plus residual tail risk

Fix a confidence level $\alpha \in (0, 1)$ and a risk-aversion or capital-conversion weight $\lambda > 0$. For each quoted offer, define

$$J(j) = \pi_j + \lambda \widehat{CVaR}_\alpha(X_{\cdot,j}), \quad X_{\cdot,j} = (X_{1,j}, \ldots, X_{n,j}). \quad (6.2)$$

The selected contract is

$$j^* \in \arg\min_{j\in\{1,\ldots,M\}} \{ \pi_j + \lambda \widehat{CVaR}_\alpha(X_{\cdot,j}) \}. \quad (6.3)$$

If there is a premium budget $B > 0$, replace the feasible set by $\{j : \pi_j \le B\}$.

Remark 6.2 *(Existence of Optimal Quoted Treaty).* Because the menu $\mathcal{M}$ is finite and non-empty, and the objective J(j) evaluates to a finite real number for each element, the minimum in (6.3) is trivially attained. The practical challenge lies not in existence, but in efficiently evaluating the empirical CVaR across the menu, which is facilitated by the convex Rockafellar–Uryasev representation.

Proposition 6.3 *(Consistency for a fixed finite-menu).* Assume $L_1, L_2, \ldots$ are i.i.d. copies of an integrable loss L, and the quoted menu $\mathcal{M}$ is fixed and finite. For each j, let $X_j = L \wedge K_j$. Then $\widehat{CVaR}_\alpha(X_{\cdot,j}) \to CVaR_\alpha(X_j)$ almost surely for each j. Consequently, the minimum objective converges almost surely, and if the population minimizer is unique, the selected empirical index is eventually equal to it almost surely.

Remark 6.3 *(MILP extension).* When treaty selection must be coupled with portfolio-level constraints such as premium budgets, capital limits, or rating-agency stress tests, the finite-menu problem can be cast as a mixed-integer linear program using the Rockafellar–Uryasev representation of CVaR. For a standalone menu of size $M < 20$, simple enumeration is computationally trivial. The MILP formulation becomes relevant when integrating treaty selection with broader enterprise risk optimization models or when selecting from a large discretised layer space, where modern solvers handle menus of several hundreds efficiently.

6.3 Simulation study with 1,000 loss scenarios

To provide a more comprehensive numerical study, consider $n = 1000$ annual aggregate-loss scenarios, measured in millions of euros. The scenarios are generated from the compound model $L = \Sigma^{N}_{r=1} Z^r + CB$, $N \sim Poisson(2)$, $Z^r \sim Lognormal(0.35, 1.05^2)$,

where $B \sim Bernoulli(0.035)$ and $C \sim Lognormal(2.7, 0.75^2)$ is an occasional catastrophe shock. All variables are independent. In the generated sample, the empirical median, 90th, 95th, and 99th percentiles of L are 3.51, 13.09, 18.31, and 36.95, respectively.

The quoted stop-loss menu is

$$(K_j, \pi_j) \in \{(2, 7.0), (5, 4.5), (10, 2.4), (20, 1.1), (40, 0.3)\},$$

again in millions of euros. Table 1 reports the retained-loss CVaR values for several confidence levels.

Table 1: Simulation-based quoted stop-loss menu.

Retention Kj	Premium πj	$\widehat{CVaR}_{\alpha.90}$	$\widehat{CVaR}_{\alpha.95}$	$\widehat{CVaR}_{\alpha.99}$
2	7.0	2.00	2.00	2.00
5	4.5	5.00	5.00	5.00
10	2.4	10.00	10.00	10.00
20	1.1	17.60	19.79	20.00
40	0.3	21.42	27.44	39.91

Note: All monetary quantities are in millions of euros; n = 1000 aggregate-loss scenarios are used. The table reports empirical CVaR of the retained loss $X = L \wedge Kj$ at confidence levels $\alpha \in \{0.90, 0.95, 0.99\}$.

Table 2 shows the sensitivity of the selected treaty to the confidence level α and the capital-conversion weight λ. The results have the expected economic interpretation. A small λ gives greater weight to quoted premium and therefore favours higher retention. A larger λ gives more weight to retained tail risk and therefore selects lower retention.

International Journal of Engineering, Science and Environment
Peer-Reviewed Open Access • CC BY 4.0 • Quarterly
Vol. 1 | Issue 1 | Apr–Jun 2026

Table 2: Sensitivity of the finite-menu objective $J(j) = \pi_j + \lambda \widehat{CVaR}_\alpha(X_{\cdot,j})$.

Confidence level α	Weight λ	Selected retention K*	Objective value J(K*)
0.90	0.10	40	2.44
0.90	0.30	10	5.40
0.90	0.60	5	7.50
0.95	0.10	40	3.04
0.95	0.30	10	5.40
0.95	0.60	5	7.50
0.99	0.10	20	3.10
0.99	0.30	10	5.40
0.99	0.60	5	7.50

Note: All retentions, premiums, CVaR values, and objective values are in millions of euros; n = 1000 scenarios are used. For λ = 0.30, the objective value is identical (5.40) for K = 10 across all confidence levels α. This occurs because, for this specific weight, the marginal reduction in the empirical CVaR exactly offsets the difference in quoted premiums between the selected retentions, resulting in a tie in the objective function.*

Finally, Table 3 gives a direct numerical check of the fixed-menu CVaR consistency statement in Proposition 6.3. For each sample size, 200 independent replications are used, except for n = 5000, where 100 replications are used. The benchmark is a Monte Carlo approximation based on 10^6 scenarios. The reported error is the maximum absolute error over the five quoted retentions.

Table 3: Finite-menu CVaR convergence check for α = 0.95 and λ = 0.30.

Scenario size n	Mean max. CVaR error	90% quantile of max. error	Correct menu selection
50	5.878	11.231	0.885
100	3.679	7.293	0.990
250	2.457	5.066	1.000
1000	1.236	2.478	1.000
5000	0.686	1.347	1.000

Note: Errors are in millions of euros and are computed relative to a 10^6-scenario benchmark. The last column reports the proportion of replications selecting the same retention as the benchmark objective.

6.4 Benchmark against a logistic GLM

To demonstrate the practical utility of the admissibility wrapper, we benchmark the proposed sigmoidal model against a standard logistic GLM [12]. We use a synthetic nonlinear probability surface to generate a binary death, default, or failure indicator:

$$P(Y = 1 \mid a, z) = \sigma(-7 + 0.085a + 0.65z + 0.9 \sin((a - 50)/7) + 0.45(z_+)^2),$$

where a is age or asset age and z is a standardised covariate. A training sample of size 5000 and an independent test sample of size 50000 are generated. A logistic GLM is compared with a one-hidden-layer sigmoidal model with 12 hidden units and a final probability link. Both models are evaluated on the same test set.

This benchmark is not intended to replace a full empirical study on proprietary insurance data, nor does it claim that neural networks should universally replace GLMs. Its purpose is to demonstrate that the proposed admissible sigmoidal class can seamlessly capture complex, non-linear covariate interactions that a standard GLM misses, while strictly preserving probability admissibility through the final sigmoid link. In practice, this framework allows actuaries to leverage the predictive power of modern ML without sacrificing the interpretability and mathematical rigor required by regulatory and pricing standards.

Table 4: Benchmark against a logistic GLM on a nonlinear probability-valued actuarial failure surface.

Model	Test log-loss	Brier score	AUC
Logistic GLM	0.4306	0.1392	0.8047
Sigmoidal one-hidden-layer model	0.4075	0.1319	0.8268

Note: Lower log-loss and Brier score are better; higher AUC is better. The sigmoidal model improves predictive accuracy while preserving probability admissibility through the final sigmoid link.

6.5 Layered extensions

If the quoted offers are finite layers with attachment K_j and limit M_j, then

$$I_j(L) = (L - K_j)_+ - (L - K_j - M_j)_+, \quad X_{i,j} = L_i - I_j(L_i).$$

The optimization is unchanged after replacing (6.1) by these retained layer losses. Additional constraints can incorporate reinsurer credit quality, reinstatement premiums, aggregate limits, capital targets, or rating-agency stress tests.

7. CONCLUSION

Sigmoidal approximations are highly effective in actuarial modeling when they are embedded in admissible architectures. A final sigmoid link gives probability-valued outputs; a positive hazard link gives monotone survival curves; bounded output links stabilise reserves and value functions; compact parameter sets support existence and consistency of empirical fitting.

Crucially, this model-agnostic wrapper allows actuaries to divorce the data-fitting capability of Machine Learning from the structural constraints of the profession, enabling the safe deployment of modern data science tools.

The most operational part of the framework is reinsurance selection from quoted offers. Stop-loss reinsurance is a call layer on aggregate losses: the insurer receives the layer above the retention, the reinsurer sells it for premium, and the insurer's retained loss is capped. By recognizing that market quotes incorporate frictions and capital costs absent in theoretical models, a finite-menu of quoted retentions can be evaluated directly on scenarios by minimizing premium plus empirical CVaR of retained losses. This creates a transparent, market-aware, and implementable connection between bounded data-driven actuarial maps and real-world tail-risk transfer.

ACKNOWLEDGEMENT

The author thanks the anonymous reviewers for their constructive comments and suggestions.

CONFLICT OF INTEREST

The author declares that he is a member of the Editorial Board of the *International Journal of Engineering, Science and Environment (IJESE)*, the journal in which this article is published. To ensure an independent and unbiased review process, the author was excluded from all editorial activities related to this manuscript, including reviewer selection, peer-review management, editorial assessment, and the final publication decision. The manuscript was handled independently by another editor and underwent the journal's standard peer-review process. The author declares that there are no other competing interests.

FUNDING STATEMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

[1] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.

[2] S. Asmussen and H. Albrecher, *Ruin Probabilities*, 2nd ed. Singapore: World Scientific, 2010.

[3] J. D. Cummins and P. Trainar, "Securitization, insurance, and reinsurance," *Journal of Risk and Insurance*, vol. 76, no. 3, pp. 463–492, 2009.

[4] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals and Systems*, vol. 2, no. 4, pp. 303–314, 1989.

[5] N. A. Doherty, *Integrated Risk Management for Financial Firms*. Upper Saddle River, NJ, USA: Prentice Hall, 2000.

[6] H. U. Gerber and E. S. W. Shiu, "On the time value of ruin," *North American Actuarial Journal*, vol. 2, no. 1, pp. 48–72, 1998.

[7] J. C. Hull, *Options, Futures, and Other Derivatives*, 10th ed. Pearson, 2018.

[8] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.

[9] R. Kaas, M. Goovaerts, J. Dhaene, and M. Denuit, *Modern Actuarial Risk Theory: Using R*, 2nd ed. Berlin, Germany: Springer, 2008.

[10] M. Leshno, V. Y. Lin, A. Pinkus, and S. Schocken, "Multilayer feedforward networks with a nonpolynomial activation function can approximate any function," *Neural Networks*, vol. 6, no. 6, pp. 861–867, 1993.

[11] A. J. McNeil, R. Frey, and P. Embrechts, *Quantitative Risk Management: Concepts, Techniques and Tools*, 2nd ed. Princeton, NJ, USA: Princeton University Press, 2015.

[12] P. McCullagh and J. A. Nelder, *Generalized Linear Models*, 2nd ed. London, U.K.: Chapman and Hall, 1989.

[13] W. K. Newey and D. McFadden, "Large sample estimation and hypothesis testing," in *Handbook of Econometrics*, vol. 4, R. F. Engle and D. L. McFadden, Eds. Amsterdam, The Netherlands: Elsevier, 1994, pp. 2111–2245.

[14] H. H. Panjer and G. E. Willmot, *Insurance Risk Models*. Schaumburg, IL, USA: Society of Actuaries, 1992.

[15] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *Journal of Risk*, vol. 2, no. 3, pp. 21–41, 2000.

[16] R. T. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1443–1471, 2002.