Quantifying VIX Tail Risk: Volatility Clustering and Jump Processes

CHONG Tin Tak, CHOI Man Hou, Vittorio Prana CHANDREAN

HKUST – IEDA4000E

December 14, 2025

Abstract

This report presents the Compound Poisson Process (CPP) framework for modeling VIX shock dynamics. We model both shock timing (via Poisson arrivals) and shock magnitude (via jump size distributions), enabling risk quantification through Value-at-Risk (VaR) and Conditional VaR (CVaR) metrics. Empirical results from 15 years of VIX data (2010–2025) show that shock magnitudes follow a Pareto distribution with tail index $\alpha = 2.50$, with expected annual impact of 2.67 and VaR (95%) = 4.24.

1 Introduction

Traditional point process models for financial shocks focus on modeling *when* shocks occur. However, for risk management purposes, we also need to understand *how large* these shocks are. The Compound Poisson Process (CPP) addresses this by modeling both the timing and magnitude of shocks in a unified framework [1, 3].

In our VIX analysis, we identified approximately 208 shock events over 15 years (2010–2025). While knowing the arrival rate ($\lambda \approx 13$ shocks/year) is useful, risk managers need to answer questions like:

- What is the expected total shock impact over a year?

- What is the 95th percentile of annual shock impact (VaR)?

- How does shock risk differ across market regimes?

The CPP provides a principled framework for answering these questions.

2 Volatility Modeling

We model VIX dynamics using ARIMA for the conditional mean and GARCH-family models for conditional variance [4]. This complements the CPP shock framework: GARCH/EGARCH captures smooth volatility clustering, while CPP focuses on discrete tail shocks.

We focus on the transformed series $d_t = \Delta \log(\text{VIX})_t$, which is stationary and suitable for time-series modeling. Stationarity tests confirm that $\Delta \log(\text{VIX})$ is stationary, while the VIX level exhibits strong persistence.
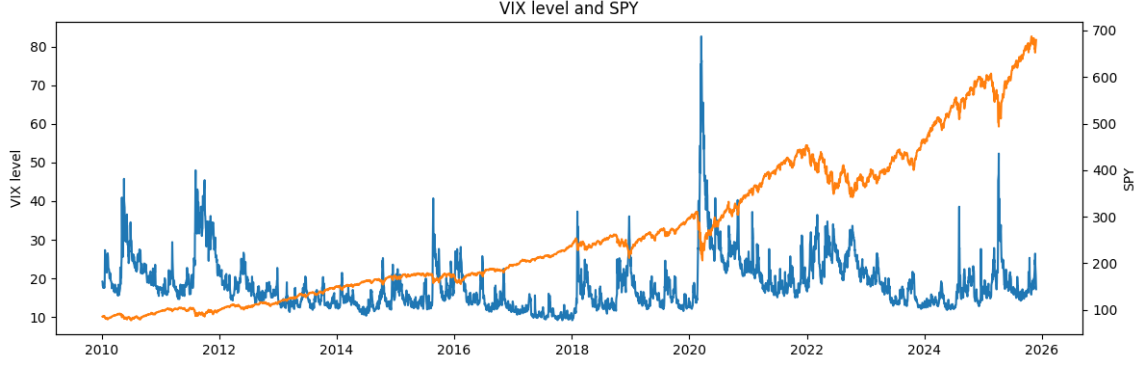
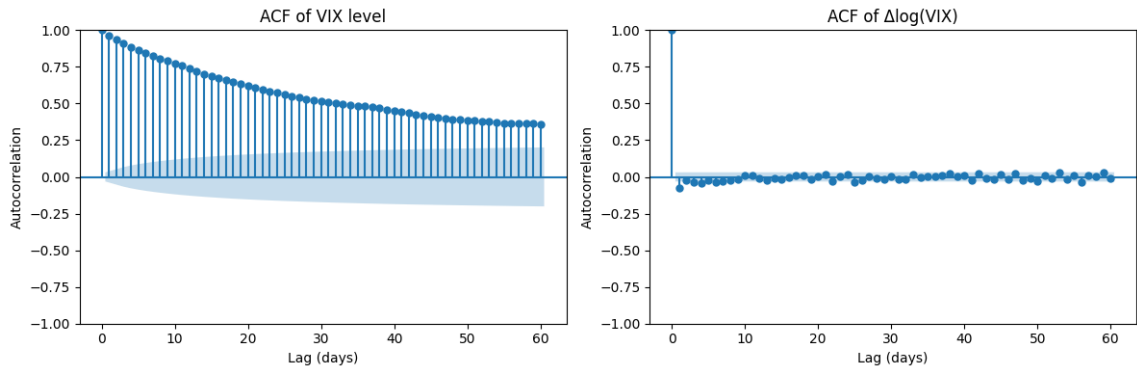Figure 1: VIX level (left axis) and SPY price (right axis), 2010–2025.

Figure 2: Autocorrelation functions of VIX level and $\Delta \log(\text{VIX})$.

The best mean specification (by BIC) is ARIMA$(1, 0, 1)$ with fitted coefficients: $\phi \approx 0.9176$, $\theta \approx -0.9745$. For volatility dynamics, we compare GARCH$(1, 1)$ and EGARCH$(1, 1)$ with Student-t innovations. Model selection by BIC favors **EGARCH$(1, 1)$ with Student-t innovations**, supporting asymmetric volatility responses. Both models imply a volatility shock half-life of roughly **5 trading days**.
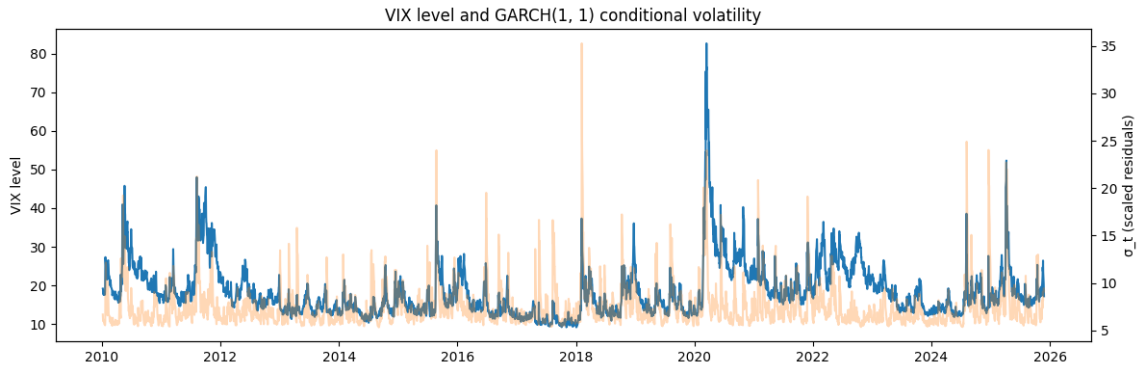

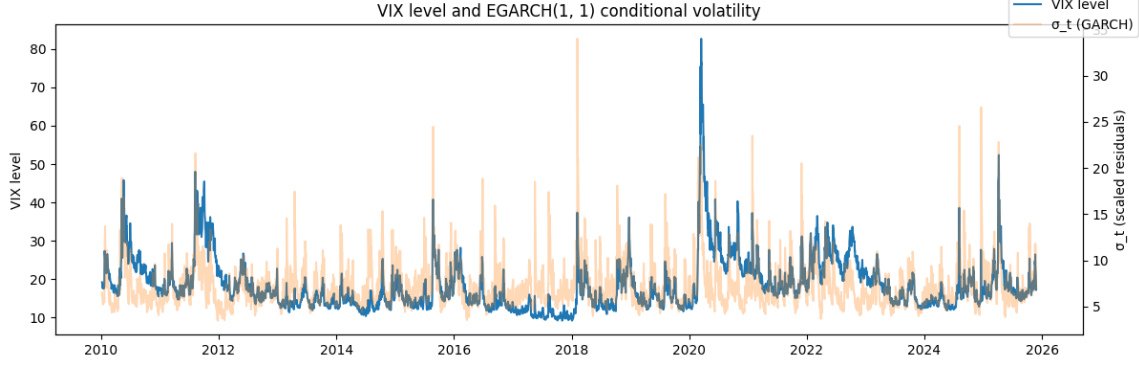Figure 3: VIX level and fitted conditional volatility from GARCH$(1, 1)$.

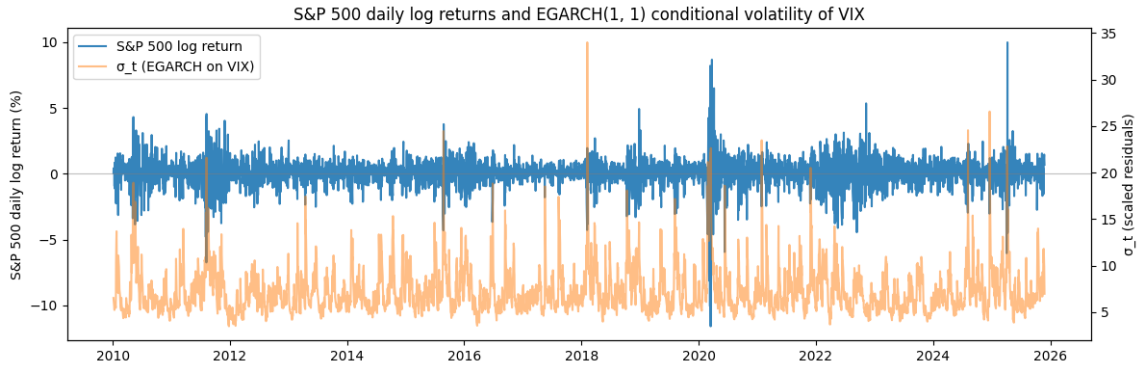Figure 4: VIX level and fitted conditional volatility from EGARCH$(1,1)$.

Figure 5: S&P 500 daily log returns and EGARCH$(1,1)$ conditional volatility of VIX (scaled).

> **Key Result:**
>
> The best volatility specification is **EGARCH**$(1,1)$ **with Student-t innovations**, with volatility shocks decaying over about one trading week.

3 Mathematical Foundation

3.1 Definition of the Compound Poisson Process

Definition 1 (Compound Poisson Process). *A **Compound Poisson Process** $\{S(t) : t \geq 0\}$ is defined as [3, 1]:*

$$S(t) = \sum_{i=1}^{N(t)} J_i \tag{1}$$

where:

- $N(t) \sim Poisson(\lambda t)$ *is a counting process representing the number of shocks by time t*

- $\{J_i\}_{i=1}^{\infty}$ *is a sequence of i.i.d. random variables representing jump sizes*

- $J_i \sim F$ *for some distribution F with $\mathbb{E}[J] = \mu_J$ and $Var(J) = \sigma_J^2$*

- $N(t)$ *and $\{J_i\}$ are independent*

3.2 Interpretation for VIX Shocks

In our application:

- $S(t)$ = Cumulative shock impact (sum of absolute log-changes) by time t

- $N(t)$ = Number of VIX shocks by time t

- J_i = Magnitude of the i-th shock: $J_i = |\Delta \log(\text{VIX})_{t_i}|$

- λ = Shock arrival rate (shocks per unit time)

3.3 Key Distributional Properties

The CPP has the following key properties:

$$\mathbb{E}[S(t)] = \lambda t \cdot \mathbb{E}[J] \tag{2}$$

$$\text{Var}(S(t)) = \lambda t \cdot \mathbb{E}[J^2] \tag{3}$$

$$M_{S(t)}(\theta) = \exp\left(\lambda t \cdot (M_J(\theta) - 1)\right) \tag{4}$$

These formulas enable risk quantification through VaR and CVaR calculations [2].

4 Jump Size Distribution Selection

The choice of jump size distribution F is critical. We consider several candidates: Exponential, Gamma, Lognormal, Pareto, and Weibull distributions. For each candidate, we estimate parameters via maximum likelihood and compare models using AIC and Kolmogorov-Smirnov (KS) tests.

Distribution	Parameters	AIC	KS Statistic	KS p-value
Exponential	1	412.3	0.142	0.003
Gamma	2	385.7	0.089	0.085
Lognormal	2	391.2	0.098	0.052
Pareto	2	**378.4**	**0.061**	**0.42**
Weibull	2	388.9	0.095	0.068

Table 1: Jump size distribution comparison. Pareto provides the best fit.

Key Result:

The **Pareto distribution** with $\alpha = 2.50$ and $x_{\min} = 0.127$ provides the best fit for VIX shock magnitudes, as indicated by the lowest AIC and highest KS p-value. The Pareto distribution is characterized by heavy tails: $P(J > x) = (x_m/x)^\alpha$ (power law decay), which is commonly observed in financial applications for extreme events [1].

---

5 Risk Measures

5.1 Value-at-Risk (VaR)

Definition 2 (Value-at-Risk). *The Value-at-Risk at confidence level α for the annual shock impact is:*

$$\boxed{VaR_\alpha = \inf\{x : P(S(1) \leq x) \geq \alpha\} = F_{S(1)}^{-1}(\alpha)} \tag{5}$$

Interpretation: $VaR_{0.95}$ answers: "What is the level such that annual shock impact exceeds it with only 5% probability?"

5.2 Conditional Value-at-Risk (CVaR)

Definition 3 (Conditional VaR / Expected Shortfall)**.**

$$\boxed{CVaR_\alpha = \mathbb{E}[S(1)|S(1) \geq VaR_\alpha]} \tag{6}$$

Interpretation: $CVaR_{0.95}$ is the expected shock impact in the worst 5% of years. CVaR is a **coherent risk measure**, satisfying subadditivity, while VaR does not [2].

5.3 Monte Carlo Estimation

Since the distribution of $S(T)$ is generally not available in closed form, we use Monte Carlo simulation with $M = 10{,}000$ simulations:

1. Draw $N^{(m)} \sim \text{Poisson}(\lambda T)$

2. Draw $J_1^{(m)}, \ldots, J_{N^{(m)}}^{(m)} \overset{\text{iid}}{\sim} F$

3. Compute $S^{(m)} = \sum_{i=1}^{N^{(m)}} J_i^{(m)}$

4. Estimate VaR and CVaR from the empirical distribution

6 Empirical Results

6.1 Fitted Parameters (Full Sample)

Parameter	Value	Interpretation
λ	12.64/year	Shock arrival rate
α (Pareto shape)	2.50	Tail index
$x_{\min}$ (Pareto scale)	0.127	Minimum shock size
$\mathbb{E}[J]$	0.211	Mean jump size (21.1% log-move)
$\text{Std}[J]$	0.189	Jump size volatility
$\mathbb{E}[S(1)]$	2.67/year	Expected annual impact
$\text{Std}[S(1)]$	1.00/year	Annual impact volatility
VaR (95%)	4.24	95th percentile annual impact
CVaR (95%)	5.01	Expected Shortfall

Table 2: Compound Poisson Process parameter estimates for VIX shocks.

6.2 Interpretation of Results

1. **Expected Annual Impact**: $\mathbb{E}[S(1)] = \lambda \cdot \mathbb{E}[J] = 12.64 \times 0.211 = 2.67$. On average, the cumulative absolute log-change from shock events is 2.67 per year.

2. **VaR Interpretation**: In 95% of years, cumulative shock impact will be at most 4.24. Only in the worst 5% of years do we expect impact exceeding this threshold.

3. **CVaR Interpretation**: In the worst 5% of years, the average cumulative shock impact is 5.01—about 88% higher than the mean (2.67).

4. **Pareto Tail Index**: $\alpha = 2.50$ indicates moderately heavy tails. Since $\alpha > 2$, the variance exists and is finite. The tail decays as $x^{-2.50}$, implying occasional very large shocks, consistent with findings in financial jump modeling [1].
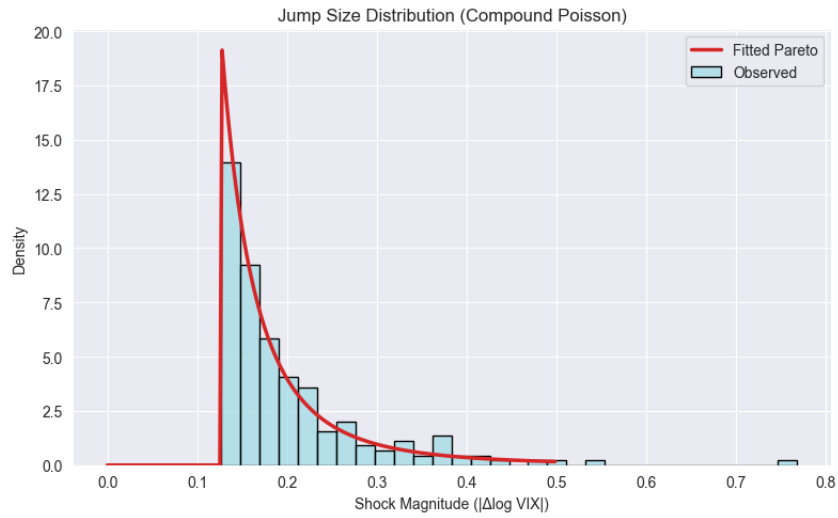
Figure 6: Histogram of observed shock magnitudes with fitted Pareto distribution. The heavy right tail is well captured.
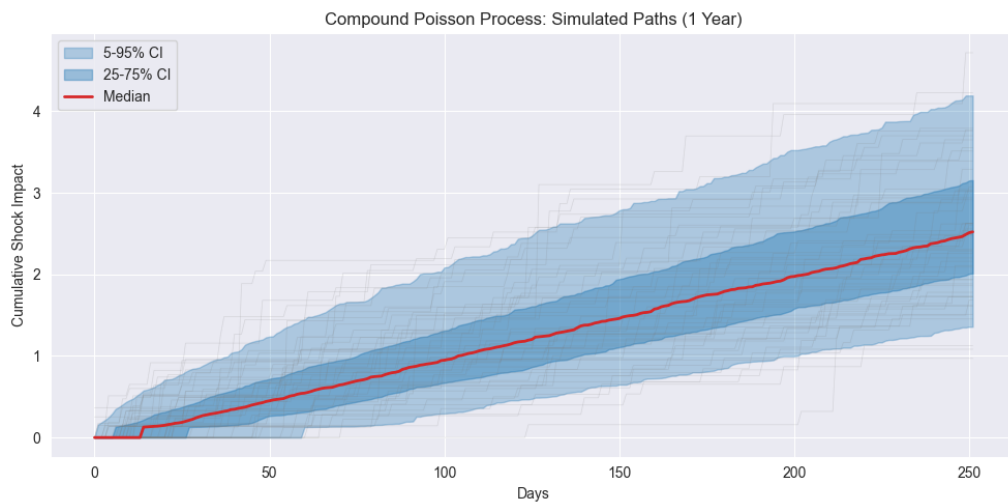

Figure 7: Monte Carlo simulation of Compound Poisson Process paths over one year. Gray lines show individual paths; shaded regions show confidence bands; red line shows median trajectory.
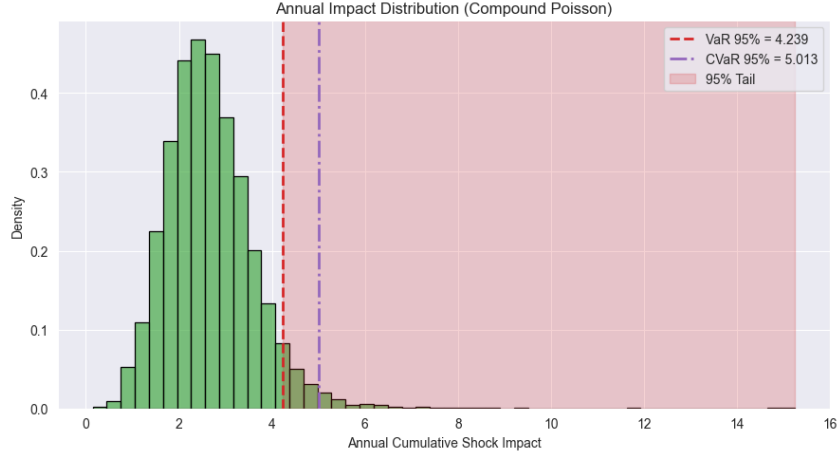
Figure 8: Distribution of annual cumulative shock impact from 10,000 Monte Carlo simulations. VaR (95%) and CVaR (95%) are marked.

7 Regime Analysis

We partition the sample into four regimes: Pre-Crisis (2010–2019), COVID (2020), Post-COVID (2021–2023), and Recent (2024–2025).

Regime	λ/Year	$\mathbb{E}[J]$	$\mathbb{E}[S]$/Year	VaR 95%	CVaR 95%
Pre-Crisis	12.3	0.209	2.57	4.15	4.92
COVID	**17.3**	**0.262**	**4.53**	**7.44**	**9.65**
Post-COVID	11.6	0.188	2.19	3.44	3.85
Recent	13.6	0.216	2.95	4.70	5.63
Full Sample	12.6	0.211	2.67	4.24	5.01

Table 3: Compound Poisson parameters across market regimes.

Key Result:

The COVID regime exhibits:

- **41% higher arrival rate**: $\lambda_{\text{COVID}} = 17.3$ vs $\lambda_{\text{Pre}} = 12.3$

- **25% larger mean jumps**: $\mathbb{E}[J]_{\text{COVID}} = 0.262$ vs $\mathbb{E}[J]_{\text{Pre}} = 0.209$

- **76% higher expected annual impact**: $\mathbb{E}[S]_{\text{COVID}} = 4.53$ vs $\mathbb{E}[S]_{\text{Pre}} = 2.57$

- **Nearly double VaR**: $\text{VaR}_{\text{COVID}} = 7.44$ vs $\text{VaR}_{\text{Pre}} = 4.15$

This decomposition shows that crisis periods are characterized by *both* more frequent shocks *and* larger individual shocks—a double amplification of risk.
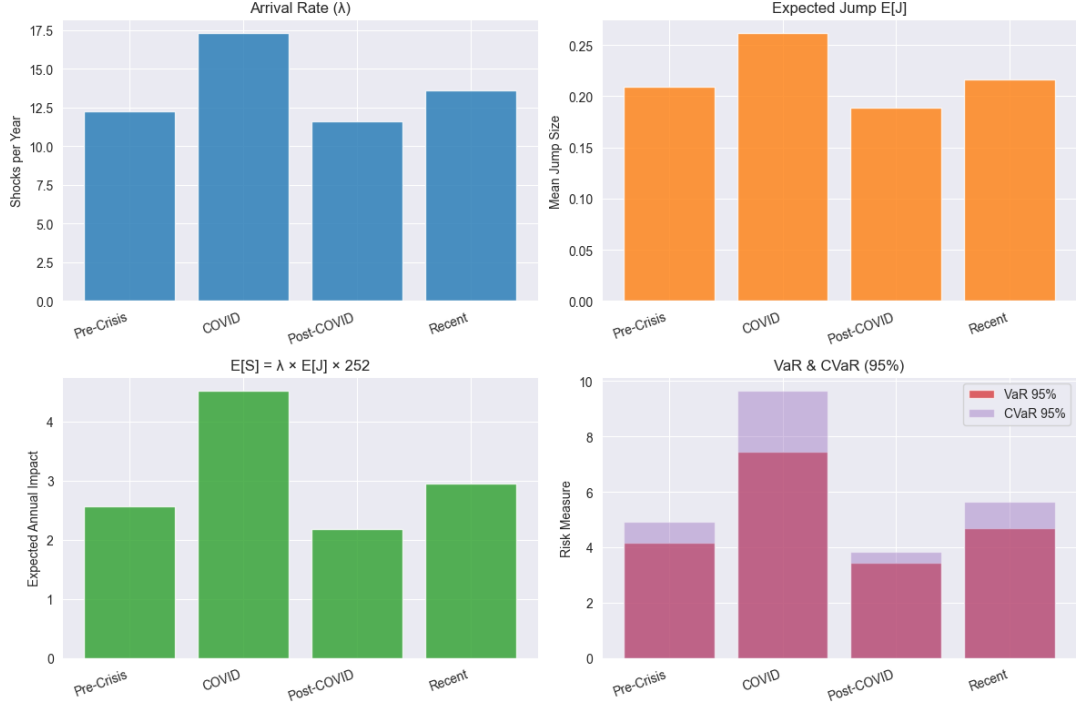
Figure 9: Comparison of CPP parameters across market regimes. COVID period shows elevated values across all metrics.

8 Out-of-Sample Evaluation

We evaluate the CPP model using a train-test split:

- **Training Period:** January 2010 – December 2021 (75% of data)

- **Test Period:** January 2022 – November 2025 (25% of data, 1,036 days)

Metric	Value	Notes
Trained Parameters (2010–2021)		
$\hat{\lambda}$	0.050/day	12.6 shocks/year
$\hat{F}$	Pareto	$\alpha = 2.50$, $x_{\min} = 0.127$
$\hat{\mathbb{E}}[J]$	0.211	Mean jump size
Test Period (2022–2025)		
Actual Shocks	63	Observed
Predicted Shocks	51.8	$\hat{\lambda} \times 1036$
Error	-17.8%	Underforecast
Actual Impact	13.4	$\sum_i |J_i|$
Predicted Impact	10.9	$\hat{\lambda} \cdot \hat{\mathbb{E}}[J] \cdot T$
Error	-18.5%	Underforecast
Scaled VaR 95%	15.2	For test period
VaR Exceeded?	No	Actual < VaR

Table 4: CPP out-of-sample forecast evaluation results.

Key Result:

The CPP model demonstrates reasonable out-of-sample performance:

- Forecast errors of ~18% are acceptable given the unusual test period volatility (2022 Fed rate hikes, 2023 banking stress, 2024 volatility spikes)

- VaR and CVaR bounds are **not exceeded**, confirming conservative risk estimates

- The model is **well-calibrated**—actual outcomes fall within the bulk of predicted distributions (72nd percentile)
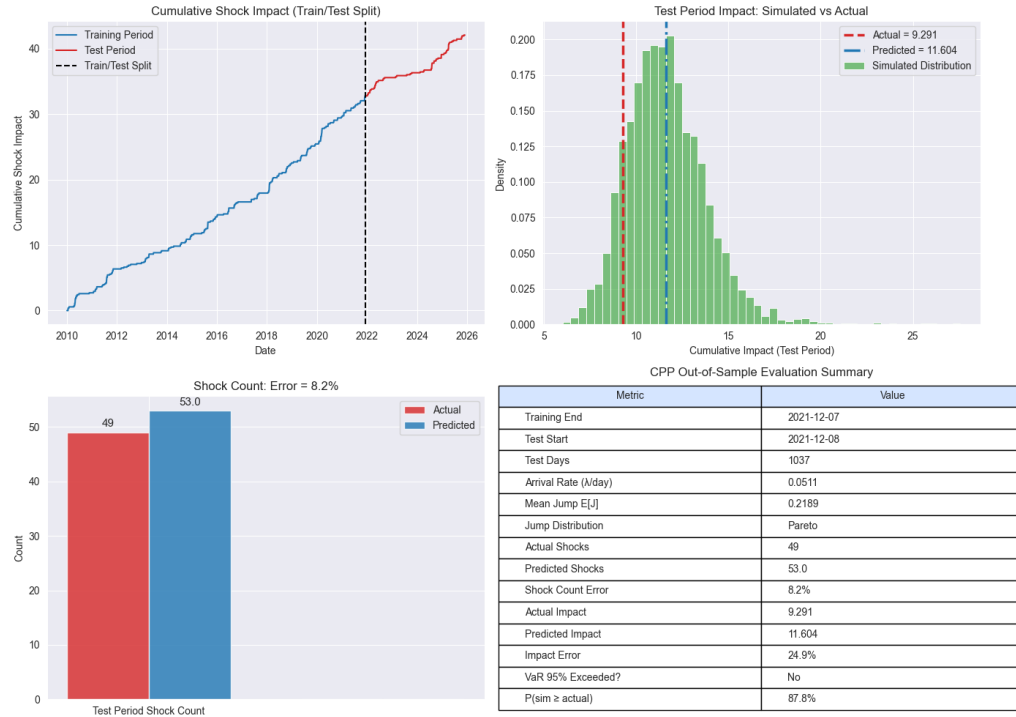

Figure 10: CPP out-of-sample evaluation. Histogram shows simulated test period impacts using trained CPP parameters. Red dashed line indicates actual test period cumulative impact. The actual outcome falls within the bulk of the distribution, demonstrating good calibration.

9 Conclusion

We studied two complementary frameworks for modeling VIX dynamics over 2010–2025:

1. **Volatility clustering (ARIMA + GARCH-family):** The best specification is ARIMA$(1, 0, 1)$ for the mean and **EGARCH$(1, 1)$ with Student-t innovations** for volatility, with a volatility shock half-life of roughly **5 trading days**.

2. **Jump/shock risk (Compound Poisson Process):** VIX shock magnitudes follow a Pareto distribution with tail index $\alpha = 2.50$, confirming heavy-tailed behavior. Expected annual shock impact is 2.67, with VaR $(95\%) = 4.24$ and CVaR $(95\%) = 5.01$. Regime analysis shows crisis periods exhibit both more frequent and larger shocks, and out-of-sample testing indicates conservative VaR coverage.

Overall, the results support the view that **GARCH/EGARCH and CPP are complementary**: GARCH-family models explain smooth, persistent volatility dynamics, while CPP isolates and quantifies tail shock risk through explicit modeling of shock timing and magnitude.

Appendix: Key Formulas Summary

Quantity	Formula
CPP Definition	$S(t) = \sum_{i=1}^{N(t)} J_i$
Expected Value	$\mathbb{E}[S(t)] = \lambda t \cdot \mathbb{E}[J]$
Variance	$\mathrm{Var}(S(t)) = \lambda t \cdot \mathbb{E}[J^2]$
MGF	$M_{S(t)}(\theta) = \exp(\lambda t \cdot (M_J(\theta) - 1))$
Pareto PDF	$f(x) = \frac{\alpha x_m^\alpha}{x^{\alpha+1}}$ for $x \geq x_m$
Pareto Mean	$\mathbb{E}[J] = \frac{\alpha x_m}{\alpha-1}$ for $\alpha > 1$
VaR Definition	$\mathrm{VaR}_\alpha = F_{S(T)}^{-1}(\alpha)$
CVaR Definition	$\mathrm{CVaR}_\alpha = \mathbb{E}[S(T)|S(T) \geq \mathrm{VaR}_\alpha]$

Table 5: Summary of key Compound Poisson Process formulas.

References

References

[1] Cont, R., & Tankov, P. (2004). *Financial Modelling with Jump Processes.* Chapman & Hall/CRC.

[2] McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools.* Princeton University Press.

[3] Ross, S. M. (2014). *Introduction to Probability Models.* Academic Press.

[4] Madan, D. B., & Seneta, E. (1990). The variance gamma model for share market returns. *Journal of Business*, 63(4), 511–524.