

Chronology-Constrained Residual Option Pricing with Local Strike-Shape Diagnostics across BTC and ETH Windows

Peilin Meng*

Zhengzhou Foreign Language School, Zhengzhou, China

**Corresponding author: Peilin Meng.*

Abstract

Deribit option trades form an irregular sequence across strikes and expiries, so random splitting can leak later market information into model fitting. We freeze four exact-expiry 24-hour BTC and ETH windows and divide each one chronologically into 70/15/15 partitions. Eight methods predict USD prices. The deterministic comparators are fitted once; the direct multilayer perceptron, unbounded residual network, and chronology-constrained residual pricer (CCRP) are run with five seeds. SVI is fitted by exact expiry using training implied volatilities only. We also inspect prices on a fixed 15-minute reference grid for elementary bounds, monotonicity, and convexity over linear moneyness. The RMSE winner changes in every window. CCRP beats its matched unbounded residual in three of four windows, but records two wins and two losses against each of Black-Scholes, previous trade, and SVI. Its bounded map produces no bound or convexity violations on the audited grids; three monotonicity violations occur in the BTC windows. The evidence supports a limited comparison under temporal controls, not stable predictive superiority or complete arbitrage consistency for asynchronous trades.

Keywords

option pricing, residual learning, chronological evaluation, strike-shape diagnostics, crypto options

1. Introduction

Public crypto-option trades do not arrive as synchronized volatility surfaces. They form an event stream spanning contracts, strikes, call/put types, and exact expiries, with long quiet periods for some instruments and clustered activity for others. A random split of that stream can place a later trade in the fitting sample and an earlier trade in the test sample. The resulting error no longer describes a causal pricing exercise.

We examine a narrower question: how do familiar statistical and machine-learning pricers behave when both the time boundary and the available information are fixed in advance? The evidence comes from four Deribit BTC and ETH asset-day windows. Every method uses the same chronological 70/15/15 partition. The comparison covers previous trade, Black-Scholes, Ridge, XGBoost, a direct MLP, training-only raw SVI, an

unbounded residual network, and CCRP. A separate reference grid records price-bound, monotonicity, and convexity failures. It is a diagnostic of the fitted price map and never feeds back into test selection.

2. Related Work

Black-Scholes supplies the analytic reference used throughout the experiments [1]. Its constant-volatility assumption is restrictive, but the price is unambiguous once the inputs are fixed and can be reproduced for every trade. Heston replaces constant volatility with stochastic variance and admits a semi-analytic price [2]. We omit it because the project does not contain an independently verified chronological calibration; adding an unstable implementation would make the comparison less, rather than more, informative.

SVI provides a second, surface-based comparator. Its compact parameterization is widely used, and later variants state precisely which static-surface conditions they impose [3, 4]. Our baseline is deliberately the raw form: it is calibrated separately for each exact expiry from training-partition implied volatilities. Validation and test IV labels are unavailable to the fit.

Neural derivative pricing predates the present deep-learning vocabulary. Hutchinson, Lo, and Poggio compared network-based pricing and hedging with classical alternatives [5]; later work surveys fast derivative pricing, hedging, and calibration with machine learning and formulates deep hedging under a risk objective [6, 7]. Recent studies also apply machine-learning regression to cryptocurrency options and explicitly model time-sequenced European-option observations [8, 9]. These settings motivate learned comparators, but they do not remove the need for an analytic anchor and a common temporal split.

Bitcoin-option studies also document market-specific information in volatility slopes and jumps, jump and convenience-yield dynamics, and smile-aware delta hedging [10, 11, 12]. Our target differs from those explanatory and hedging questions: we reconstruct held-out transaction prices from a deliberately small feature set. The literature therefore motivates the crypto-option setting without serving as evidence that any one method should dominate the four windows.

Cohen, Reisinger, and Wang consider a stronger object: neural-SDE dynamics whose state space preserves linear no-arbitrage inequalities across an option book [13]. That dynamic surface model is not a drop-in baseline for the present trade stream. CCRP only bounds a point prediction and adds local strike-shape penalties. We therefore report remaining failures directly and do not attribute the arbitrage scope of the neural-SDE construction to our map.

The split rule follows the same caution required for dependent sequences: serial dependence and nonstationarity can invalidate an ordinary random-split interpretation [14]. Within each asset-day window, trades are ordered first by observed time and then by trade identifier. Training, validation, and test blocks are cut only after that ordering is fixed.

3. Method

Figure 1 separates the common data boundary from model construction and post hoc diagnostics. All methods receive the same chronological partitions. The two residual routes use the same Black-Scholes anchor, which is added only when the final price is formed and is not an input to either residual network. The unbounded route forms an unconstrained anchored price, whereas CCRP passes the anchored residual through the bounded price map and is regularized by local moneyness checks. Predictions from Black-Scholes, SVI, the unbounded residual, and CCRP enter the terminal reference-grid audit; the audit has no path back to fitting, validation, or checkpoint selection.

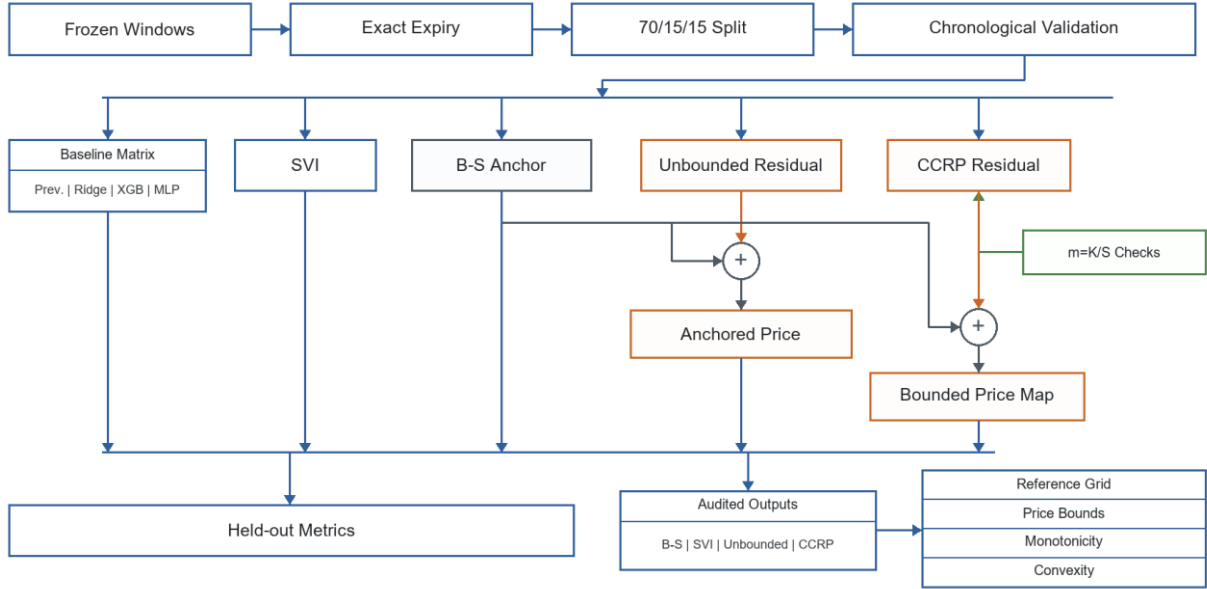
3.1 Data Boundary and Features

The frozen sample contains Deribit BTC and ETH trades from June 29 and July 13, 2026, giving two 24-hour windows per asset. A record carries the option type, trade and expiration timestamps supplied by the exchange, strike, contemporaneous index price, and the latest available historical-volatility observation. Endpoint definitions and request fields follow the exchange's public API documentation [15]. Learned models receive only option type, log moneyness, time to the recorded expiration, and historical volatility. Trade

implied volatility is reserved for SVI. The analytic calculations use a zero risk-free rate as a disclosed convention; no rate is inferred from the transactions.

Ties in exchange time are resolved by trade identifier before the ordered records are cut into 70/15/15 training-validation-test partitions.

Figure 1: Chronology-constrained evaluation pipeline. The frozen temporal partitions feed every comparator. Black-Scholes, SVI, the unbounded residual, and CCRP are inspected on the same terminal reference grid, which does not feed back into fitting or checkpoint selection. Only CCRP uses the bounded price map and local moneyness regularization



3.2 Compared Models

All eight comparators return an option price in USD: causal previous trade, zero-rate Black-Scholes, Ridge, XGBoost, direct MLP, raw SVI, an unbounded residual MLP, and CCRP. At the test boundary, previous trade starts with the training and validation history. It may then use a test trade only after that trade has occurred and only for the same contract; otherwise it falls back to Black-Scholes. The deterministic models are fitted once. Seeds 13, 29, 47, 71, and 101 are used for the direct MLP and both residual networks. Among these models, only CCRP combines the bounded map with a shape penalty.

Ridge uses training-standardized features augmented with squared log-moneyness and squared maturity; its penalty is selected from 10^{-6} , 10^{-4} , 10^{-2} , and 1 by validation MSE. XGBoost uses 240 trees, a learning rate of 0.05, row and column sampling ratios of 0.9 and 1.0, histogram tree construction, one worker, and the first registered seed; depth 3 or 5 is chosen by validation MSE. The direct MLP has widths 4-64-64-1, SiLU hidden activations, and a Softplus price output. It is trained with AdamW, learning rate 0.001, weight decay 10^{-5} , batches of at most 512, an 80-epoch cap, patience 12, and the five registered seeds.

3.3 Anchored Bounded Residual

Here S denotes the index price observed with the trade, K denotes the strike, and P denotes the USD option price. We work with linear moneyness $m=K/S$ and relative price $q=P/S$; c denotes the option type. A zero-rate Black-Scholes price computed from prior historical volatility gives the relative anchor q_- . The elementary admissible interval is

$$(l, u) = \begin{cases} (\max(1-m, 0), 1), & c = 1 \\ (\max(m-1, 0), m), & c = 0 \end{cases} \quad (1)$$

Here $c=1$ denotes a call and $c=0$ a put; l and u are the corresponding lower and upper relative-price bounds. The residual input is the standardized vector $x=(c, \log m, \tau, \nu)$, where τ is time to expiration and ν is prior historical volatility. It passes through a two-layer SiLU network. Zero initialization of the output layer makes the initial residual zero, so the starting prediction is the analytic anchor. CCRP maps the result back into the interval:

$$\hat{q} = l + (u - l) \sigma \left[\text{logit} \left(\frac{q_{BS} - l}{u - l} \right) + R_{\theta}(x) \right] \quad (2)$$

where the normalized anchor inside the logit is clipped to the open unit interval for numerical stability.

3.4 Training Objective

For a mini-batch, training minimizes

$$L(\theta; B) = \frac{1}{|B|} \sum_{i \in B} (\hat{q}_i - q_i)^2 + \lambda_{\text{shape}} L_{\text{shape}} \quad (3)$$

where $\lambda_{\text{shape}} = 0.001$. For compact notation, $f_{\theta}(m, z)$ denotes the relative-price prediction at linear moneyness m , while the network itself receives $\log m$; z denotes the remaining feature tuple of option type, maturity, and historical volatility. For each i in B , the code sets $h_i = \max(10^{-4}, 0.01 m_i)$ and evaluates $\hat{q}_i^- = f_{\theta}(m_i - h_i, z_i)$, $\hat{q}_i^0 = f_{\theta}(m_i, z_i)$, and $\hat{q}_i^+ = f_{\theta}(m_i + h_i, z_i)$. With $s_i = 1$ for a call and $s_i = -1$ for a put, the one-sided monotonicity penalty is

$$L_{\text{mono}} = \frac{1}{|B|} \sum_{i \in B} \left[\max(0, s_i (\hat{q}_i^+ - \hat{q}_i^-)) \right]^2 \quad (4)$$

The symmetric local second difference and its convexity penalty are

$$D_i^{(2)} = \frac{\hat{q}_i^+ - 2\hat{q}_i^0 + \hat{q}_i^-}{h_i^2} \quad (5)$$

Both terms operate in linear moneyness and are local regularizers; they do not impose the full set of static-arbitrage conditions.

For a matched comparison, the unbounded residual uses the same features, two-layer SiLU architecture, zero-initialized output, optimizer, checkpoint rule, seeds, learning rate of 0.001, 80-epoch cap, and patience of 12. Removing the bounded map and shape term is the only structural change.

3.5 Training-only Svi

The SVI comparator uses the raw total-variance form

$$w(k) = a + b \left[\rho(k - m_{\text{SVI}}) + \sqrt{(k - m_{\text{SVI}})^2 + \sigma_{\text{SVI}}^2} \right] \quad (6)$$

where a denotes the variance offset, b denotes the scale, ρ denotes skew, m_{SVI} denotes location, and σ_{SVI} denotes width. Here K again denotes strike, F denotes the forward, and S denotes the contemporaneous index price; $k = \log(K/F)$ and $F=S$ under the zero-rate convention. Training uses implied volatility only from the training partition of each exact expiry and minimizes unweighted total-variance error. A fit requires at least five strikes. The initial vector is $(0.01, 0.2, 0, 0, 0.2)$, with a in $[0, 4]$, b in $[10^{-5}, 5]$, ρ in $[-0.999, 0.999]$, m_{SVI} in $[-2, 2]$, and σ_{SVI} in $[10^{-5}, 3]$. One deterministic start and at most 800 function evaluations are used. The fitted variance is extrapolated without clipping and converted to a USD option price; a missing or failed exact-expiry fit falls back to Black-Scholes.

3.6 Reference-Grid Diagnostics

The primary audit uses 15-minute test buckets. Eligible slices are grouped by time bucket, exact expiration timestamp, and option type, and must contain at least three distinct strikes. Within a slice, the reference time, index price, and historical volatility are their respective medians. Slices with nonpositive reference spot or maturity are discarded, and the remaining distinct strikes are sorted before model evaluation. This fixes one common information tuple within each slice rather than treating asynchronous trades as a contemporaneous market surface.

For each eligible slice, prices are ordered as $m_1 < \dots < m_J$. Unlike the symmetric training perturbation, the observed strike grid is generally unequal. For adjacent grid prices p_j , the audit computes

$$d_j = \frac{p_{j+1} - p_j}{m_{j+1} - m_j}; \Delta_j^{\text{ug}} = d_j - d_{j-1} \quad (7)$$

Calls require d_{j0} , puts require d_{j0} , and convexity requires Δ_{j0} . Because spot is fixed within a grid slice, using strike instead of $m=K/S$ multiplies every slope by a positive constant and leaves these signs unchanged. Black-Scholes calculations use float64 arithmetic and absolute and relative tolerances of 10^{-10} . Every count is reported as violations over eligible checks. Black-Scholes and SVI have no failures. CCRP has no bound or convexity failures and monotonicity counts of 2/240, 1/246, 0/20, and 0/46. For the same four windows, the unbounded residual records 33/327, 7/333, 7/29, and 6/65 bound failures and 29/240, 5/246, 2/20, and 0/46 monotonicity failures; its convexity counts are zero. These checks describe the model on the chosen grid, not the full asynchronous market.

4. Experiments

The test bed consists of BTC and ETH trades from two UTC days, June 29 and July 13, 2026. These four asset-day windows are processed independently. Expiration timestamps come directly from the exchange rather than being reconstructed from rounded time-to-expiry fields. After ordering, the first 70 % of trades in each window go to training, the next 15 % to validation, and the final 15 % form the held-out test set.

Table 1: Window-level pricing errors in USD. Entries are RMSE / MAE; parentheses give RMSE seed SD

Method	BTC Jun. 29	BTC Jul. 13	ETH Jun. 29	ETH Jul. 13
Previous trade	137.79 / 28.61	118.77 / 27.65	9.56 / 3.45	9.84 / 2.39
Black-Scholes	544.76 / 325.55	380.09 / 166.66	16.13 / 9.11	11.21 / 3.37
Ridge	1829.71 / 1327.03	1375.09 / 710.50	257.31 / 208.53	57.96 / 21.38
XGBoost	596.24 / 149.98	498.52 / 206.72	71.02 / 16.17	25.79 / 8.80
Direct MLP	1986.36 (23.63) / 1017.24	1524.77 (95.71) / 772.17	127.65 (0.18) / 45.02	62.45 (2.17) / 22.31
SVI	217.23 / 160.12	188.28 / 109.69	6.60 / 3.24	8.96 / 3.94
Unbounded residual	1122.48 (517.84) / 398.20	90.92 (9.85) / 64.22	399.70 (281.64) / 320.90	3.98 (0.15) / 1.84
CCRP	644.34 (669.75) / 208.92	75.62 (2.09) / 43.67	40.97 (6.59) / 19.17	5.67 (1.17) / 2.43

Table 2: Data characteristics of the four frozen asset-day windows

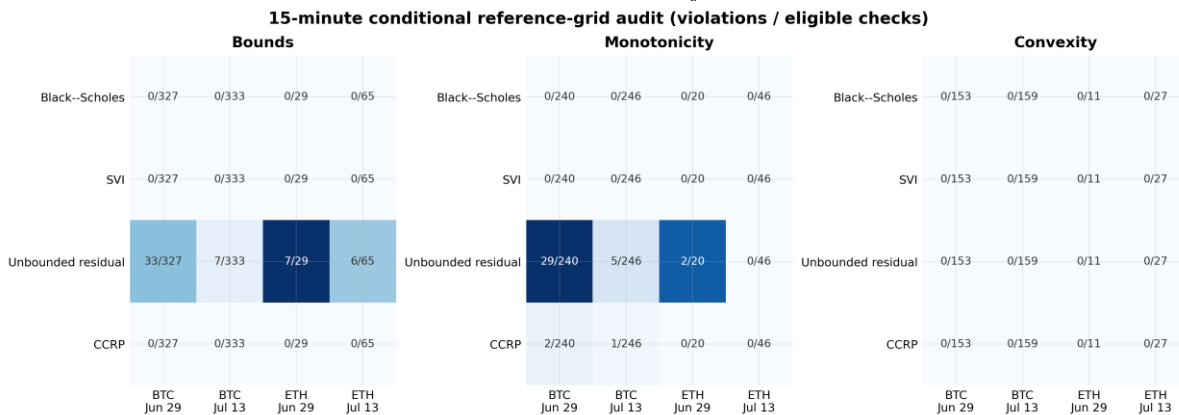
Asset-day	Trades	Instruments	Expiries	Strikes	Active min.
BTC Jun. 29	11,993	407	13	72	967
BTC Jul. 13	10,378	418	12	69	1,242
ETH Jun. 29	2,105	258	13	53	592
ETH Jul. 13	3,444	263	12	56	913

The raw Deribit requests, including active and expired instrument responses, historical-volatility responses, deduplicated trade identifiers, exclusion reasons, and split manifests, are versioned together. Trade implied volatility serves diagnostic purposes only and is never a learned feature. The risk-free rate is fixed at zero in the analytic reference.

Eight methods are evaluated in the same output unit, USD option price: causal previous trade with Black-Scholes fallback, zero-rate Black-Scholes, Ridge, XGBoost, direct MLP, raw SVI, an unbounded Black-

Scholes residual MLP, and CCRP. The previous-trade comparator uses only strictly earlier same-contract test trades, falling back to the contemporaneous Black-Scholes reference where no earlier transaction is available. SVI is fit separately for each exact expiry using training-set implied volatilities only; validation and test labels are excluded.

Figure 2: Reference-grid comparison of violations and eligible-check counts for Black-Scholes, SVI, the unbounded residual, and CCRP across the four windows



For direct MLP, unbounded residual, and CCRP, Table I reports the mean RMSE over seeds 13, 29, 47, 71, and 101 together with its seed standard deviation. The deterministic rows contain a single fit. MAE is shown beside RMSE so that a few large pricing errors can be distinguished from a broad shift in absolute error.

Figure 2 narrows the grid comparison to Black-Scholes, SVI, the unbounded residual, and CCRP. Using float64 Black-Scholes arithmetic and 10^{-10} absolute and relative tolerances, CCRP records bound counts of 0/327, 0/333, 0/29, and 0/65. Its monotonicity counts are 2/240, 1/246, 0/20, and 0/46, and its convexity counts are 0/153, 0/159, 0/11, and 0/27. Black-Scholes and SVI remain at zero for every eligible check. The unbounded residual crosses price bounds and violates monotonicity, although no convexity failure is observed. XGBoost and direct-MLP counts are kept in the reproduction package. None of these grid counts is a complete asynchronous-market test.

The four rows in Table II differ materially in trading activity. A window contains 258-418 instruments, 12-13 exact expiries, 53-72 strikes, and 592-1,242 active minutes. We keep the windows separate because pooling these distinct event streams would hide both their time ordering and their different contract coverage.

Window winners are previous trade on BTC June 29, CCRP on BTC July 13, SVI on ETH June 29, and the unbounded residual on ETH July 13. CCRP records two wins and two losses against Black-Scholes, previous trade, and SVI; three wins and one loss against XGBoost and the matched unbounded residual; and four wins against Ridge and direct MLP. These counts are descriptive only and do not imply statistical significance. Previous-trade hit rates are 98.83\

5. Discussion and Limitations

The four windows produce different winners rather than a single ranking. Previous-trade hit rates are 98.83 %, 97.21 %, 94.67 %, and 93.35 % for the four respective windows, computed over samples where an earlier same-contract trade exists.

The bounded map enforces the implemented pointwise call and puts intervals. It does not enforce calendar consistency, put-call parity, bid-ask executability, or hedging performance, and the fixed reference grid is not a complete asynchronous-market arbitrage test. The audit therefore measures local model geometry at disclosed reference inputs; it is not a profitability or deployment result.

The evidence is limited to four single-exchange 24-hour transaction windows, a zero-rate convention, and a compact feature set. Transaction targets can reflect bid-ask effects, stale trades, tick sizes, and short-lived order flow. A verified chronological Heston calibration was not completed, so Heston remains a missing

stochastic-volatility comparator rather than evidence in favor of CCRP. No population-level inference is drawn from the four date-asset units.

6. Reproducibility

Reproduction starts from the exact-expiry request inventory and the frozen response hashes. Code, environment specifications, exclusion and deduplication rules, split manifests, seeds, model configurations, predictions, metrics, resource logs, and request integrity hashes will be released upon acceptance. Together these records trace a reported number to its observation set and preprocessing choices. Raw responses subject to provider restrictions will not be redistributed. Running the released code against the frozen responses reconstructs the computation; a later query to Deribit should be treated as a different market sample.

7. Conclusion

No method dominates these four windows: previous trade, CCRP, SVI, and the unbounded residual each obtain one RMSE win. CCRP removes all observed price-bound and convexity failures on the fixed grids, but three BTC monotonicity failures remain. The result is therefore more specific than an arbitrage-free pricing claim. It shows what the bounded residual map enforces under the stated inputs and where local shape errors still appear.

The comparison is also small. It covers four 24-hour windows from one exchange, assumes a zero rate, and uses a compact feature set. SVI is fitted without validation or test IV labels, whereas a verified chronological Heston calibration is absent. Wider date and exchange coverage would be needed before drawing a stable ranking. Calendar consistency, put-call parity, executable bid-ask prices, hedging error, and profitability remain outside the evidence reported here.

References

- [1] F. Black and M. Scholes, "The pricing of options and corporate liabilities," *Journal of Political Economy*, vol. 81, no. 3, pp. 637-654, 1973.
- [2] S. L. Heston, "A closed-form solution for options with stochastic volatility with applications to bond and currency options," *The Review of Financial Studies*, vol. 6, no. 2, pp. 327-343, 1993.
- [3] J. Gatheral and A. Jacquier, "Arbitrage-free SVI volatility surfaces," *Quantitative Finance*, vol. 14, no. 1, pp. 59-71, 2014.
- [4] G. Guo, A. Jacquier, C. Martini, and L. Neufcourt, "Generalized arbitrage-free SVI volatility surfaces," *SIAM Journal on Financial Mathematics*, vol. 7, no. 1, pp. 619-641, 2016.
- [5] J. M. Hutchinson, A. W. Lo, and T. Poggio, "A nonparametric approach to pricing and hedging derivative securities via learning networks," *The Journal of Finance*, vol. 49, no. 3, pp. 851-889, 1994.
- [6] J. De Spiegeleer, D. B. Madan, S. Reyners, and W. Schoutens, "Machine learning for quantitative finance: Fast derivative pricing, hedging and fitting," *Quantitative Finance*, vol. 18, no. 10, pp. 1635-1643, 2018.
- [7] H. Buehler, L. Gonon, J. Teichmann, and B. Wood, "Deep hedging," *Quantitative Finance*, vol. 19, no. 8, pp. 1271-1291, 2019.
- [8] A. Brini and J. Lenz, "Pricing cryptocurrency options with machine learning regression for handling market volatility," *Economic Modelling*, vol. 136, Art. no. 106752, 2024.
- [9] L. Liang and X. Cai, "Time-sequencing European options and pricing with deep learning: Analyzing based on interpretable ALE method," *Expert Systems with Applications*, vol. 187, Art. no. 115951, 2022.
- [10] T. Chen, J. Deng, and J. Nie, "Implied volatility slopes and jumps in Bitcoin options market," *Operations Research Letters*, vol. 55, Art. no. 107135, 2024.
- [11] J. E. Hilliard and J. T. D. Ngo, "Bitcoin: Jumps, convenience yields, and option prices," *Quantitative Finance*, vol. 22, no. 11, pp. 2079-2091, 2022.

- [12] C. Alexander and A. Imeraj, “Delta hedging Bitcoin options with a smile,” *Quantitative Finance*, vol. 23, no. 5, pp. 799-817, 2023.
- [13] S. N. Cohen, C. Reisinger, and S. Wang, “Arbitrage-free neural-SDE market models,” *Applied Mathematical Finance*, vol. 30, no. 1, pp. 1-46, 2023.
- [14] C. Bergmeir, R. J. Hyndman, and B. Koo, “A note on the validity of cross-validation for evaluating autoregressive time series prediction,” *Computational Statistics & Data Analysis*, vol. 120, pp. 70-83, 2018.
- [15] Deribit, “Deribit API documentation: Public market-data endpoints,” 2026. [Online]. Available: <https://docs.deribit.com/>. Accessed: Jul. 15, 2026.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Open Access

This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

