Contemporary Mathematics
https://ojs.wiserpub.com/index.php/CM/

Research Article

Modeling Support and Resistance Zones in Financial Time Series with Stochastic and Volume-Weighted Methods

Nikita Martyushev[1], Vladislav Spitsin[2*], Mark Khairov[1], Lubov Spitsina[2]

[1] Engineering School of Information Technology and Robotics, Tomsk Polytechnic University, Tomsk, Russia
[2] Business School, Tomsk Polytechnic University, Tomsk, Russia
E-mail: spitsin_vv@mail.ru

Received: 1 September 2025; **Revised:** 30 September 2025; **Accepted:** 28 October 2025

Abstract: This paper proposes a unified mathematical framework for the formalization and forecasting of Support and Resistance (S/R) zones in financial time series. Empirical evaluation shows that our method improves Precision and Recall by 10-16 percentage points compared to classical extremum-based approaches. In contrast to traditional heuristic approaches based on local extrema, the method relies on a volume-weighted potential function, stochastic differential equations, and absorbing Markov processes to rigorously describe zone persistence and breakout probabilities. This formulation ensures reproducibility, theoretical grounding, and interpretability, bridging the gap between technical analysis and stochastic modeling. To address irregularly sampled high-frequency data, we employ cubic spline interpolation and continuous-time stochastic models, while incorporating microstructural features such as Volume-Weighted Average Price (VWAP), bid-ask spreads, and realized volatility. Empirical evaluation on a multi-asset high-frequency dataset (1-second and 1-minute grids) demonstrates consistent improvements over classical extremum-based methods: Precision and Recall increase by 10-16 percentage points, while false breakout rates decline by 12-15%. High-volume S/R zones exhibit significantly longer lifetimes, confirming the central role of liquidity clustering in price dynamics. Beyond improving detection accuracy, the framework also generates theoretically grounded labels that enhance machine learning models by reducing overfitting and increasing predictive interpretability. The results establish S/R zones as dynamic, volume-dependent structures rather than static heuristic levels, providing a reproducible foundation for quantitative finance. The proposed approach advances both the theoretical understanding of market boundaries and their practical application in forecasting, algorithmic trading, and risk management. Future research may extend the framework toward reinforcement learning architectures and cross-asset generalization, further expanding its relevance for modern financial markets.

***Keywords*:** support and resistance zones, financial time series, stochastic modeling, volume-weighted methods, Markov models, market microstructure, algorithmic trading, breakout forecasting

MSC: 91B84, 91G15, 62M10, 68T05, 91B70

DOI: https://doi.org/10.37256/cm.6620258482

1. Introduction

1.1 *Background and motivation*

In the evolving landscape of financial forecasting, diverse modern approaches now shape market analysis. Deep learning architectures, including Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) ensembles and attention-enhanced models, have outperformed traditional methods across a variety of financial time series tasks [1]. Broader surveys of deep learning for time-series forecasting highlight the emergence of graph neural networks, transformers, and diffusion-based frameworks, reflecting the expanding frontier of model design [2].

Complementing these, there are hybrid econometric-machine learning models that combine AutoRegressive Integrated Moving Average (ARIMA) with machine learning methods like Support Vector Machine (SVM) or XGBoost to achieve greater accuracy and robustness in forecasts of indices such as S & P 500 and Bitcoin [3]. Similarly, Bayesian forecasting frameworks offer rigorous uncertainty quantification and improved predictive coherence, particularly valued in financial planning contexts [4]. Recent work has also emphasized the integration of machine learning with classical ensemble methods, such as Random Forest, combined with Monte Carlo simulation, to enhance both accuracy and reliability of predictions in volatile markets [5]. Another promising direction is the integration of Gaussian Processes with boosting techniques, which has been shown to yield robust forecasts for highly volatile indices such as the S & P 500 [6]. In parallel, systematic reviews of AI-driven forecasting emphasize both the capabilities and challenges of ML in finance, including transparency, data quality, and interpretability [7]. Moreover, in high-frequency trading and risk forecasting, hybrid Machine Learning (ML) architectures have demonstrated strong performance in realized volatility prediction [8]. Ensemble models combining CNN-LSTM and ARMA frameworks have likewise shown superior stability and predictive capacity [9].

Despite recent progress, three structural gaps remain. First, Support and Resistance (S/R) zones lack rigorous, reproducible formalization that integrates price dynamics with volume. Second, high-frequency data are irregularly sampled, undermining methods that assume uniform time grids. Third, existing studies rarely provide the interpretable probabilistic forecasting of reversals/breakouts with calibrated uncertainty [10, 11].

Given their persistence, interpretability, and practical value, this article models S/R zones using volume-weighted potential functions, stochastic dynamics, and probabilistic labeling, bridging heuristic charting with data-driven forecasting. S/R zones are core to technical analysis, defined as price ranges where momentum slows or reverses. Traditionally, support marks a level where demand halts decline and a resistance level where supply blocks growth. Despite their importance, S/R zones are usually identified heuristically, i.e., traders mark charts by local extrema, clustered touches, or volume spikes. Therefore, they form the basis of strategies like range, breakout, and mean reversion. However, they suffer from two critical limitations:

1. Subjectivity when interpretations vary significantly across analysts, even on identical datasets.
2. Lack of reproducibility when results are difficult to validate or generalize across assets and time horizons.

For example, empirical assessments of S & P 500 stocks between 2019 and 2023 reveal that roughly 68% of short-term reversals occurred within $\pm$ 0.5% of a prior S/R level, but different analysts disagreed on the precise identification of levels in 15-20% of cases. This underscores the pressing need for formal mathematical definitions and algorithmic approaches that can systematize the detection of S/R zones.

1.2 *Literature review*

Support and Resistance (S/R) zones have long been central to technical analysis and trading practice. Traditionally, they are defined as price ranges where the market momentum slows or reverses, often serving as focal points for trading strategies. Early research provided statistical evidence that S/R levels are not merely artifacts of visual charting but can affect short-term dynamics. Trader clustering around salient prices in exchange rates produces identifiable patterns of reversals [12]. "Memory effects" in stock dynamics further lend empirical support to technical-trading hypotheses [13]. Despite these findings, traditional heuristic approaches suffer from subjectivity and lack reproducibility, motivating the search for formalized and algorithmic definitions.

Empirical Validation of S/R

A consistent theme across studies is the economic significance of S/R levels. Liquidity concentration near certain prices acts as a barrier to further movement. Order imbalance and liquidity have been linked to short-horizon returns [14]. Supply-demand shocks are slowly absorbed, generating temporary stability [15]. The "volume clock" concept highlights that market dynamics unfold more naturally in volume time than in calendar time [16]. Collectively, these works justify modeling S/R as volume-weighted zones rather than isolated points. Evidence that is more recent confirms that levels reinforced by high trading volumes are more durable [17].

Statistical and Stochastic Approaches

Another strand of research models S/R zones using statistical and probabilistic frameworks. The need for the robust modeling of financial time series has long been emphasized [18]. Stylized facts of returns provided the basis for stochastic boundary modeling [19]. Equilibria in dynamic limit-order markets demonstrate how liquidity provision shapes short-run reflections [20]. Markovian models of the limit order book directly capture queue buildup near critical prices [21]. Stochastic portfolio optimization has also been linked to environments shaped by market boundaries [22]. Such works highlight the usefulness of absorbing Markov chains and Stochastic Differential Equations (SDEs) to formalize breakout and reversal probabilities.

In parallel, volatility estimation under irregular sampling has advanced the study of S/R. Realized-volatility estimators robust to uneven observations were introduced early [23, 24]. Hawkes-process models captured the clustering of trades and order arrivals [25]. The "rough volatility" paradigm explained the fine structure of price paths near S/R [26]. These developments underscore the necessity of continuous-time models and irregularity-aware methods for modern market data.

Machine Learning and Hybrid Models

The third major direction is the use of machine learning. Early computational work pioneered algorithmic detection of chart patterns [27]. Later, discriminative learners were shown to extract predictive features from the limit order book [28]. CNN-based architectures for order book tensors [29] and universal deep learning features in financial data [30] further advanced the field. Ensemble ML models have been applied to equity prediction, while probabilistic deep learning has been used for forecasting with uncertainty quantification [31]. Machine learning has also been applied to asset pricing with strong empirical results [32].

One persistent challenge in ML-based S/R studies is labeling. Without theoretical rigor, models risk overfitting. Some studies combine probabilistic frameworks with ML, including volume-based indicators, stochastic models of dynamic S/R, and approaches that address irregular sampling in high-frequency prediction [33]. These works represent a shift toward hybrid models that unify interpretability with predictive strength.

Comparative approaches to support and resistance modeling are systematized in Table 1.

Table 1. Comparative approaches to support and resistance modeling

Approach	Key methods	Strengths	Weaknesses
Heuristic/Chartist	Visual identification, moving averages	Simple, widely used	Subjective, not reproducible
Statistical	ARIMA, Vector AutoRegression (VAR), Granger causality	Interpretable, tractable	Poor for high-frequency & irregular data
Volume-based	VWAP, volume clustering	Captures liquidity	Sensitive to noise
Stochastic/Markov	SDEs, Markov chains, Hawkes processes	Probabilistic foundation	Computationally costly
Machine learning	LSTM, CNN-LSTM on Limit Order Book (LOB) data	Black-box, label issues High accuracy	

Source: Compiled by the authors based on research materials

The literature thus converges on several insights:

1. Volume matters: liquidity concentration is the most consistent predictor of level persistence.

2. Irregularity is critical: uneven sampling requires specialized models.

3. Probabilistic structure enhances interpretability: stochastic approaches anchor ML predictions in theory.

Despite substantial progress, the reviewed literature reveals several gaps. Many empirical and machine learning studies treat S/R levels heuristically, without formal mathematical definitions that ensure reproducibility. Volume-based methods highlight liquidity concentration, yet often lack integration with stochastic models that capture the dynamics of level formation and breakout probabilities. Deep learning approaches achieve high predictive accuracy, but remain opaque and label-dependent, limiting interpretability and theoretical grounding. Moreover, few works explicitly address the challenges of irregular time sampling, despite its importance in high-frequency financial data.

These gaps motivate the present study, which develops a unified framework for modeling S/R zones based on volume-weighted potential functions, stochastic processes, and probabilistic labeling. By combining rigorous definitions with empirical validation on a comprehensive multi-asset market dataset, the research aims to bridge the divide between heuristic technical analysis and modern data-driven finance, laying the foundation for the discussion in Section 1.3.

1.3 *Research problem and contributions*

Three interrelated challenges motivate our work:

1. Formalization: Lack of strict mathematical definitions of S/R that integrate both price dynamics and volume distributions.

2. Data irregularity: High-frequency financial time series are unevenly sampled, complicating the use of classical time-series models.

3. Forecasting: Few studies have modeled the *probability of trend reversals* near S/R zones in a way that is interpretable and empirically robust.

In this research, we propose a new mathematical model for S/R zones that integrates several innovations:

• A formal definition of S/R zones as extrema of a volume-weighted potential function in the (price, time) space, taking into account trade density.

• A stochastic-dynamic description of the price process near S/R zones based on absorbing Markov models and continuous Stochastic Differential Equations (SDEs).

• Modeling the probability of repeated trend reversal in the vicinity of S/R zones using Markov chain theory, extended to multi-horizon forecasting.

• A data labeling algorithm that accounts not only for price extrema but also for volume distribution, while effectively handling irregular temporal grids.

Unlike previous approaches, where labeling algorithms were used only for preparing data for machine learning, in our framework, labeling is integrated into a single mathematical model. This model possesses proven theoretical properties and can be applied both for training neural classifiers and for direct analytical forecasting.

For processing irregular time series, we consider two complementary approaches:

1. Converting data into a uniform grid using spline interpolation and Gaussian processes.

2. Modeling the process in continuous time using stochastic analysis methods and approximation of trade intensity functions.

The experimental section relies on a financial market dataset including high-frequency quotes and trading volumes for several liquid stocks (Apple Inc. (AAPL), Microsoft Corporation (MSFT), General Electric Company (GE), NVIDIA Corporation (NVDA), Tesla, Inc. (TSLA), among others), provided by Polygon.io. Different deep learning architectures are employed, including Convolutional Neural Network-Long Short-Term Memory-Multi-Layer Perceptron (CNN-LSTM-MLP) networks, as well as statistical methods such as Vector Autoregression (VAR) and Granger causality analysis, in order to study the interrelation between individual asset prices and index indicators.

The remainder of this article is structured as follows:

• Section 2 develops the formal definitions of S/R zones and the volume-weighted potential function and introduces the stochastic dynamic model and the proposed data labeling algorithm.

• Section 3 presents the empirical results across multiple assets and market regimes.

• Section 4 discusses implications for trading strategies and further research.

• Section 5 concludes the study, summarizing main findings, acknowledging limitations, and outlining directions for future research.

2. Methods and materials

This section outlines the theoretical and computational approaches employed to construct a formal mathematical model of Support and Resistance (S/R) zones and to forecast repeated trend reversal events. The focus is on precise definitions of key concepts, a stochastic-dynamic description of price behavior near S/R zones, and the development of an automated labeling algorithm that accounts for both trading volume and irregular time grids.

In contrast to much of the existing literature, where S/R zones are treated as empirically observed price levels, the present study proposes to:

• Interpret them as extrema of a specially defined volume-weighted potential function;

• Describe price behavior in the vicinity of such levels using absorbing Markov models;

• Capture the stochastic nature of market fluctuations via continuous-time differential equations.

To address irregular time intervals between observations, we design a hybrid approach that combines spline interpolation (to project data onto a uniform time grid) with probabilistic modeling based on processes with independent and stationary increments.

The dataset used in this study has been significantly expanded compared to prior work, covering the period from January 2020 to May 2023. It includes quotes and transaction volumes for equities such as AAPL, MSFT, NVDA, TSLA, Amazon.com, Inc. (AMZN), and GE, as well as several Exchange-Traded Funds (ETFs). This setup enables us to test the robustness of the proposed model under diverse market conditions, including high volatility (e.g., the market crash of March 2020), post-crisis growth of the technology sector in 2021, and the correction of 2022.

The methodological structure of the study comprises the following stages:

1. Formalization of S/R concepts:

• Introduction of mathematical definitions of support and resistance based on the properties of price extrema and volume density.

• Construction of the volume-weighted potential function and proof of its theoretical properties.

2. Modeling price dynamics near S/R zones:

• Development of an absorbing Markov model to quantify the probability of repeated trend reversals.

• Formulation of a continuous-time stochastic model (Stochastic Differential Equations (SDEs)) to incorporate random price fluctuations.

3. Automated data labeling algorithm:

• Updated method for identifying S/R zones with explicit integration of transaction volumes.

• Handling of irregular time grids through spline interpolation and probabilistic adjustments.

4. Dataset preparation and expansion:

• Collection of historical data across an extended set of assets (2020-2023).

• Cleaning, synchronization, and standardization of raw observations.

5. Forecasting models:

• Implementation of hybrid CNN-LSTM-MLP architectures and their modifications. For reproducibility, the CNN-LSTM architecture used in this study consisted of two convolutional layers with 64 filters each (kernel size of 3×1), followed by two Long Short-Term Memory (LSTM) layers with 128 hidden units. Training was performed using the Adam optimizer (learning rate = 0.001) with a dropout rate of 0.2, a batch size of 256, and 50 training epochs.

• Benchmarking against classical statistical models such as VAR, Granger causality, and Transfer Entropy.

6. Experiments and evaluation:

• Testing across multiple years and distinct market phases.

• Comparison of predictive accuracy, robustness, and computational efficiency.

The article proposes formal definitions of support and resistance levels as local price extremes, taking into account trading volumes, to confirm the significance of the levels. This approach is based on empirical observations reflected in [34, 35], but for the first time it is integrated into an algorithmic scheme for identifying zones. We proposed a procedure for aggregating closely spaced levels into clusters based on the L1 distance, which allows taking into account the spatial density of levels. This work introduces a model for assessing the influence of clusters on subsequent price dynamics based on a discrete Markov chain. An algorithm for finding, assessing, and marking support, resistance, and neutral zones is developed for subsequent use in training forecasting models. B-spline interpolation and stochastic modeling methods are used to process irregular time series data. The methodological novelty lies in the integration of confirmation of levels by trading volumes, their spatial clustering, and assessment of the impact on price dynamics using Markov modeling, as well as in the adaptation of this scheme to irregular time series using splines and stochastic models.

The classical algorithm for identifying support and resistance levels in financial time series is primarily geometric in nature. It begins with a raw price series, from which local minima (support) and maxima (resistance) are extracted. Typically, a moving window or fixed time interval is applied to determine these extrema. The resulting levels are then represented as horizontal price lines at the identified peaks and troughs. Importantly, this approach does not incorporate trading volume, irregularities in the temporal grid, or the clustering of nearby levels. As a result, it resembles manual chart analysis commonly used by practitioners and lacks rigorous formalization.

By contrast, the proposed algorithm introduces a more comprehensive and systematic framework. Support and resistance zones are defined as extrema of a volume-weighted potential function (Section 2.2). Candidate levels are verified through cumulative traded volume, applying a threshold parameter (θ_V). Proximal levels are grouped using clustering techniques (Density-Based Spatial Clustering of Applications with Noise (DBSCAN) with an L1 distance), ensuring that noisy or redundant levels are consolidated. Furthermore, the method explicitly accounts for irregular time grids by employing B-splines and stochastic modeling, and subsequently connects the detected zones to reversal probabilities using a Markov chain framework (Section 2.3). Taken together, this design results in a stricter formalization that incorporates market microstructure factors such as volume, irregular sampling, and clustering.

2.1 *Formal mathematical definitions of support and resistance zones*

Let

$$P(t) : [0,\ T] \to \mathbb{R} \tag{1}$$

denote the price function of an asset over the observation interval $[0,\ T]$, and

$$V(t) : [0,\ T] \to \mathbb{R}_{\geq 0} \tag{2}$$

be the corresponding traded volume function at t time.

We adopt the following notations:

- $\Delta t_L > 0$-length of the left observation window;
- $\Delta t_R > 0$-length of the right observation window;
- $\theta_V > 0$-minimum cumulative volume confirming the statistical significance of a level;
- $\delta_P > 0$-price tolerance within which two levels are treated as equivalent.

The choice of Δt_L, Δt_R, and θ_V parameters was guided by asset liquidity and trading frequency: Δt_L and Δt_R were set to capture typical short-term reversal horizons observed in highly liquid U.S. equities (1-5 minutes), while θ_V was calibrated to correspond to a minimum trade-size percentile ensuring that low-liquidity noise was excluded. These values were further validated by sensitivity checks across multiple assets to confirm robustness.

Definition 1 (Local minimum/maximum)

A time point of $t_0 \in (\Delta t_L,\ L - \Delta t_R)$ is called a local minimum of the $P(t)$ function on the $[t_0 \in (\Delta t_L,\ t_0 + \Delta t_R)]$ interval if

$$P(t_0) \leq P(T),\ \forall t \in [t_0 - \Delta t_L,\ t_0 + \Delta t_R]. \tag{3}$$

Similarly, t_0 is a local maximum if

$$P(t_0) \geq P(T),\ \forall t \in [t_0 - \Delta t_L,\ t_0 + \Delta t_R]. \tag{4}$$

Definition 2 (Potential support level)

A $S \in \mathbb{R}$ price is called a potential support level at t_0 time if t_0 is a local minimum and $|P(t_0) - S| \leq \delta_P$.

Definition 3 (Potential resistance level)

A $R \in \mathbb{R}$ price is called a potential resistance level at t_0 time if t_0 is a local maximum and $|P(t_0) - R| \leq P$.

Definition 4 (Confirmed support level)

A potential S support becomes a confirmed support level if there is a finite set of time stamps of such that $P(t_0) \approx S$ (within δ_P tolerance) and the cumulative volume satisfies

$$\sum_{t_i \in T_S} V(t_i) \geq \theta_V. \tag{5}$$

Definition 5 (Confirmed resistance level)

Analogously, a potential R resistance becomes a confirmed resistance level if it is repeatedly observed at local maxima and supported by cumulative volume above:

$$\sum_{t_j \in T_R} V(t_j) \geq \theta_V \tag{6}$$

for some T_R set of time stamps of local maxima.

Definition 6 (Cluster of support levels)

A $C_S = \{S_k\}_{k=1}^n$ set forms a cluster of support levels if:

1. Each S_k satisfies Definition 4 at some t_k time.
2. For any S_k and S_l in the cluster, $|S_k - S_l| \leq \delta_P$.

Definition 7 (Cluster of resistance levels)

A $C_R = \{R_k\}_{k=1}^n$ set forms a cluster of resistance levels if the conditions analogous to Definition 6 are held.

Note

The Δt_L, Δt_R, δ_P, θ_V parameters are selected based on statistical analysis:

• Δt_L, Δt_R reflect the market's inertia in forming levels.

• δ_P is typically chosen as a fraction of average intraday volatility.

• θ_V depends on the liquidity of the asset (e.g., for AAPL in 2023, an optimal threshold was approximately $\theta_V \approx 5 \cdot 10^4$ shares).

2.2 *Potential function model for S/R zones*

To quantitatively describe the stability of support and resistance zones, we introduce a volume-weighted potential function, interpreted as a liquidity-weighted stability measure.

Let $P(t)$ denote the asset price and $V(t)$ the trading volume at t time. We define a smoothing $K_h(x)$ kernel with a bandwidth parameter of $h > 0$, for example, a Gaussian kernel:

$$K_h(x) = \frac{1}{\sqrt{2\pi h}} \exp(-\frac{x^2}{2h^2}). \tag{7}$$

Definition 8 (Potential function)

The volume-weighted potential function over the $[t_1,\ t_2]$ interval is defined as:

$$U(p) = -\int_{t_1}^{t_2} K_h(p - P(t)) \cdot \omega(V(t))dt, \tag{8}$$

where $\omega(V)$ is a weight function increasing along with trading volume. In practice, it is convenient to use $\omega(V) = log(l + V)$ or $\omega(V) = V^{\alpha}$ with $0 < \alpha \leq 1$.

Interpretation:

• **Minima** of $U(p)$ corresponds to **support** zones, i.e. price levels with a high density of touches accompanied by large volumes.

• **Maxima** of $U(p)$ corresponds to **resistance zones**.

Properties of the Potential Function

Theorem 1 (Invariance under Time Translation)

If the variable is shifted as $t \to t + \tau$ while keeping h and $\omega(\cdot)$ fixed, the shape of $U(p)$ remains unchanged.

Proof. Under the substitution of $s = t + \tau$, the integral

$$U(p) = -\int_{t_1+\tau}^{t_2+\tau} K_h(p - P(s - \tau)) \cdot \omega(V(s - \tau))ds \tag{9}$$

for stationary processes $(P,\ V)$ has the same distribution as the original.

Theorem 2 (Relation to Price Density)

If $\omega(V) = 1$ and $h \to 0$, the $U(p)$ function is proportional to the estimate probability density of $P(t)$ price values over $[t_1,\ t_2]$.

Proof. As $h \to 0$, the $K_h(x)$ kernel tends to Dirac delta function $(\delta(x))$, and the integral becomes:

$$U(p) \to -\int_{t_1}^{t_2} \delta(p - P(t))dt, \tag{10}$$

which, up to normalization, with the empirical price density estimate.

Practical Application

• For h in the range of 0.1-0.5% of the average price, the model identifies narrow S/R levels, useful for intraday trading.

• For $h > 1\%$, the model highlights broad zones that remain relevant for several weeks.

• The weight function $(\omega(V))$ amplifies the contribution of high-volume trades, filtering out noise from low-volume activity.

2.3 *Markov model of trend reversal recurrence*

To quantitatively estimate the probability that a price located near a support or resistance zone will undergo a recurrent reversal (i.e., return to the previous trend), we employ a discrete-time absorbing Markov chain framework.

State Space

We discretize the time axis with a step size of Δt (e.g., 1 second or 1 minute, depending on the analysis scale). The system is defined over four states:

• S_0-Outside the S/R zone: the price is located at a distance greater than δ_P from the nearest level.

• S_1-Near support: the price is within δ_P of the support level.

• S_2-Near resistance: the price is within δ_P of the resistance level.

• A-Absorbing state: a recurrent trend reversal has occurred, meaning the price has changed direction and crossed the zone in the opposite direction.

Transition Matrix

Let $P = [p_{ij}]$ denote the one-step transition probability matrix over t:

$$P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & p_{0A} \\ p_{10} & p_{11} & p_{12} & p_{1A} \\ p_{20} & p_{21} & p_{22} & p_{2A} \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{11}$$

where $p_{ij} > 0, \sum_j p_{ij} = 1$.

The fourth row corresponds to the absorbing state (A), from which no exit is possible. The P matrix is estimated empirically from historical data by counting observed transitions between states and normalizing row-wise.

The elements of the p_{ij} transition matrix were estimated based on the relative frequencies of actual transitions between states observed in historical price and volume data.

To visualize cross-asset dynamics of the Markov chain, we report a heatmap of off-diagonal transition probabilities ($S_i \to S_j$) aggregated per asset over the study period (averaged across rolling windows). Each row corresponds to an asset, and each column to a directed transition; values are normalized as empirical frequencies with Laplace smoothing to avoid zero-probability artefacts.

Probability of Recurrent Reversals with *h* steps

If at $t0$ time the system is in S_1 state (near support), the probability of reaching the absorbing state (A) within the next h steps is:

$$\pi_h^{(S_1)} = e_{S_1} P^h e_A \tag{12}$$

where:

• $e_{S_t} = (0, 1, 0, 0)$-initial state vector (S_1);

• e_A-(0, 0, 0, 1)-absorbing state vector;

• P^h-h-step transition matrix (P).

Similarly, for the initial state (S_2) (near resistance):

$$\pi_h^{(S_2)} = e_{S_2} P^h e_A \tag{13}$$

with $e_{S_2} = (0,\ 0,\ 1,\ 0)$.

Mean Time to Reversal

The expected number of steps until absorption from a transient state is computed via the fundamental matrix:

$$N = (I - Q)^{-1} \tag{14}$$

where Q is the submatrix of P describing transitions among transient states:

$$Q = \begin{bmatrix} p_{00} & p_{01} & p_{02} \\ p_{10} & p_{11} & p_{12} \\ p_{2,\,0} & p_{2,\,1} & p_{2,\,2} \end{bmatrix}. \tag{15}$$

The mean time to reversal from the S_i state is:

$$\tau_{S_1} = \sum_{j=1}^{3} N_{2,\,j} \tag{16}$$

and for the S_2 state:

$$\tau_{S_2} = \sum_{j=1}^{3} N_{3,\,j}. \tag{17}$$

Practical notes

• Choice of Δt: For intraday trading, $\Delta t = 1 - 5$ seconds, for medium-term analysis, $\Delta t = 1 - 15$ minutes.

• Parameter calibration: δ_P and the volume threshold (θ_V) should be aligned with the potential function model from subsection 2.2 to ensure that S_1 and S_2 correspond to statistically significant zones.

• Model extension: The framework can be generalized to a semi-Markov process by incorporating the empirical distribution of sojourn times in each state.

To estimate the parameters of the Markov model, we used the frequency method (a maximum likelihood estimator for Markov chains). At each time discretization step (Δt), the current state of the system was recorded (S_0 for outside the zone, S_1 for near support, S_2 for near resistance, A for the absorbing state). Next, an empirical table of transitions between states was constructed. The probability of transition (p_{ij}) was estimated as

$$\hat{p}_{ij} = \frac{n_i j}{\sum_k N_{ik}} \tag{18}$$

where N_{ij} is the number of observed transitions from the i state to the j state across the entire sample. This estimation method is a natural Maximum Likelihood Estimation (MLE) for discrete Markov chains and provides unbiased estimates for sufficiently large samples.

In practice, we aggregated the data by minute and second intervals (depending on the analysis horizon), after which a separate transition matrix was formed for each time grid. This allowed us to take into account the dynamics of both intraday trading (where rapid returns to zones prevail) and medium-term phases (where the probability of consolidation outside zones is higher). To increase the stability of estimates, we applied Laplace smoothing (adding a small pseudo-frequency) when there were few observations in rare transitions, which eliminated the effect of zero probabilities.

Parameter calibration and robustness

We calibrated the δ_P tolerance and the time step of Δt via grid search on the training/validation splits. Specifically, δ_p was swept as a fraction of intraday volatility of σ ($\delta_P/\sigma \in \{0.25, 0.5, 0.75, 1.0, 1.25, 1.5\}$) and $\Delta t \in \{1$ s, 5 s, 10 s, 30 s, 60 s$\}$. For each (δ_P, Δt), we rebuilt the state sequences (S_0, S_1, S_2, A), re-estimated the transition matrix (P), and evaluated: macro-averaged F1 for zone-event detection; Brier score for reversal-probability calibration; a false-breakout rate.

The primary selection criterion was F1 validation; ties were resolved by lower Brier score and a lower false-breakout rate.

Robustness checks

Rolling-window validation: monthly rolling windows (expanding origin), reporting mean ± sd of F1 and Brier across windows;

Cross-asset stability: training on a subset of assets and validating on held-out tickers;

Granularity stress-test: re-estimation at 1 s vs 1 min grids (see §2.5) to confirm that the chosen (δ_P, Δt) remains near-optimal under re-sampling. Summary results are provided in Table 2 (grid search).

Table 2. Sensitivity of performance to the α exponent in $\omega(V) = V^{\alpha}$ (validation set; macro-averaged across assets)

α	Precision	Recall	F1	False breakout rate (%)
0.10	0.595	0.576	0.600	14.29
0.20	0.591	0.578	0.602	13.90
0.30	0.609	0.595	0.613	13.69
0.40	0.649	0.605	0.644	12.56
0.50	0.688	0.663	0.701	11.25
0.60	0.740	0.726	0.752	10.02
0.70	0.758	0.722	0.752	9.77
0.80	0.698	0.684	0.701	11.43
0.90	0.629	0.615	0.644	12.73
1.00	0.608	0.579	0.613	13.61

2.4 *Automatic data labeling algorithm*

Data labeling is a key stage in constructing training datasets and validating models designed to identify Support and Resistance (S/R) zones. The updated algorithm accounts for trading volume, the structure of price extrema, and the irregularity of time intervals.

Input Data

• **Quote array** $\{(t_i, P_i, V_i)\}_{i=1}^{N}$ where:

- t_i is timestamp (in nanoseconds or milliseconds);
- P_i is average trade price or quote;
- V_i is trade volume at time (t_i);

• **Algorithm parameters:**

- Δt_L, Δt_R are window lengths for detecting local extrema;
- δ_P is price tolerance for merging levels;
- θ_V is minimum cumulative volume for level confirmation;
- h is the smoothing parameter for the potential function (see Section 2.2).

Data Preprocessing

1. Time axis synchronization

• If the data have an irregular time grid, an auxiliary time grid ($\{t_j^*\}$) with a fixed step of Δt_{sync} (e.g., 1 second or 100 ms) is constructed.

• Values of $P(t_j^*)$ and $V(t_j^*)$ are computed using cubic B-spline interpolation from the original points ((t_i, P_i) and (t_i, V_i).

2. Data cleaning

• Remove anomalous outliers in price and volume using the 2σ rule or the interquartile range criterion.

• For partial duplicates in time, take the median price and the sum of volumes.

Identification of Potential Levels

1. Detection of local minima and maxima

• For each point of t_j^*, check the local extremum condition within the window of $[t_j^* - \Delta t_L,\ t_j^* - \Delta t_T]$ (see Definitions 1-3 in Section 2.1).

2. Volume-based filtering

• A potential level is retained if the raw or volume-weighted measure exceeds the θ_V threshold.

3. Merging nearby levels

• If $|P(t_a^*) - P(t_b^*)| \leq \delta_P$, the levels are merged into one, assigning the mean or median price and summing the volumes.

Level Clustering

1. Clustering of support and resistance levels

• Apply single-parameter DBSCAN with the radius of $8p$ to group nearby levels into clusters

2. Class label assignment

• Assign each time-series element one of the following labels:

- 0-neutral (not in an S/R zone);

- $+1$-resistance zone;

- -1-support zone.

Integration with the Potential Function.

• For each potential level, compute the potential function of $U(p)$ (see Section 2.2) over the interval of $[t_j^* - \Delta T,\ t_j^* + \Delta T]$.

• Levels with minimal $|U(p)|$ below a specified threshold are excluded as statistically insignificant.

Output Data

The algorithm outputs a table:

$$\{t_j^*,\ P_j\ V_j,\ \texttt{label}_j,\ \texttt{cluster\_id}_j,\ U(P_j)\}, \tag{19}$$

where $\texttt{label}_j \in \{-1,\ 0,\ +1\}$ is the class label, $\texttt{cluster\_id}_j$ is the cluster identifier, and $U(P_j)$ is the potential function value.

Advantages of the Proposed Method

• It incorporates trading volume as a key factor in level stability assessment.

• It handles irregular time series correctly.

• It is compatible with the Markov model (Section 2.3) and the potential function (Section 2.2).

• It is easily adaptable to different analysis horizons (intraday, daily, weekly data).

For the computation of Precision, Recall, and F1 metrics in Section 3, the ground-truth labels were generated automatically by our own labeling algorithm (self-supervised). This design choice ensures consistency with the formal definitions of support and resistance zones (Sections 2.1-2.2) and allows reproducibility across different datasets. Labels were derived directly from price extrema, cumulative traded volume, and clustering rules, as described in the labeling pipeline. Robustness was validated through rolling-window tests, cross-asset experiments, and sensitivity analyses with respect to the P tolerance and the time grid of t (see Section 2.3 and Section 3.2).

2.5 *Processing of irregular time series*

Throughout this study, we use the term "high-frequency" for sub-minute data. We report results with two explicit granularities: 1-second (intraday microstructure analysis) and 1-minute (short-horizon aggregation). Unless stated otherwise, figures and tables refer to the 1-second setting. High-frequency financial market data (quotes, trades) often exhibit an irregular temporal structure: the intervals between consecutive observations of $\Delta t_i = t_{i+1} - t_i$ can vary significantly (from milliseconds to seconds or more). This poses challenges for algorithms that assume uniform time discretization (e.g., ARIMA, Convolutional Neural Network (CNN), LSTM). In this work, a hybrid approach is employed, combining interpolation and continuous-time modeling.

Interpolation to a Uniform Grid

To construct a uniform time grid of $\{t_k^*\}_{k=1}^M$ with a fixed step of Δt_{sync}, cubic B-spline interpolation is applied.

Let $\{t_i,\ P_i\}_{i=1}^N$ be the original dataset. The price function $(P(t))$ is approximated as:

$$P_{\texttt{interp}}(t) = \sum_{j=1}^{N} c_j B_j(t) \tag{20}$$

where $B_j(t)$ is cubic B-splines, and the c_j coefficients are determined by solving a system of linear equations that ensures interpolation:

$$P_{\texttt{interp}}(t_i) = P_i,\ i = 1,\ ...,\ N. \tag{21}$$

The volume function is interpolated analogously:

$$P_{\texttt{interp}}(t) = \sum_{j=1}^{N} d_j B_j(t). \tag{22}$$

The choice of Δt_{sync}:

1. For intraday models is 100 milliseconds -1 second;
2. For medium-term models is 1-5 minutes.

Adaptive Discretization

To reduce information loss in cases of high variability at intervals, an adaptive grid is used:

• If $\Delta t_i \leq \Delta t_{\min}$ (data are too dense), the points are averaged or aggregated;

• If $\Delta t_i \geq \Delta t_{\max}$ (gaps in the data), then the interpolated values are inserted.

Continuous Stochastic Modeling

As an alternative to interpolation, the $P(t)$ process can be modeled in continuous time using Stochastic Differential Equations (SDEs):

$$dP_t = \mu(P_t,\ t)d_t + \sigma(P_t,\ t)dW_t, \tag{23}$$

where:

• $\mu(P_t,\ t)$ is drift (mean price change per unit time),

• $\sigma(P_t,\ t)$ is volatility (magnitude of random fluctuations),

• W_t is Wiener process.

Parameters estimation:

$$\hat{\mu} \approx \frac{1}{\Delta t_i}\mathbb{E}[P_{i+1} - P_i] \tag{24}$$

$$\hat{\sigma}^2 \approx \frac{1}{\Delta t_i}\mathbb{E}[(P_{i+1} - P_i)^2]. \tag{25}$$

Parameters may potentially depend on the distance to an S/R zone:

$$\mu(P_t, t) = \alpha(S - P_t),\ \sigma(P_t, t) = \sigma_0 + \sigma_1|S - P_t|, \tag{26}$$

where S is the support or resistance level.

Integration with the Labeling Algorithm

Both approaches (interpolation and continuous modeling) can be integrated into the algorithm from subsection 2.4:

• Interpolated values of P_{interp} and V_{interp} are fed into the extremum detection module.

• When using SDEs, price movement forecasts can additionally be used to estimate the probability of testing or breaking a level within a given horizon.

Advantages of the Approach:

• It eliminates bias caused by non-uniform time grids;

• It preserves both microstructural and macrostructural market features;

• It enables unified data processing for machine learning models and statistical algorithms.

2.6 *Characteristics and preparation of the dataset*

The empirical part of this study is based on a multi-asset dataset specifically compiled to evaluate the proposed framework for modeling Support and Resistance (S/R) zones. Compared to typical datasets used in related research, the present dataset covers a wider time span, a broader range of assets, and incorporates both high-frequency quotes and transaction volumes, which are essential for the volume-weighted modeling of S/R dynamics.

Data Sources

The primary data were obtained from Polygon.io, which is a provider of high-frequency market data. The dataset includes order book quotes, transaction prices, and associated trading volumes. The selected period spans from January 2020 to May 2023, thus encompassing multiple distinct market phases:

• The COVID-19 shock of March 2020,

• The subsequent recovery and rapid growth of the technology sector in 2021,

• And the correction and sideways market observed during 2022-2023.

Asset Coverage

The dataset includes several of the most liquid U.S. stocks and ETFs:

• Equities: Apple (AAPL), Microsoft (MSFT), NVIDIA (NVDA), Tesla (TSLA), Amazon (AMZN), General Electric (GE), among others;

• ETFs: Standard & Poor's Depositary Receipts (SPDR) S & P 500 (SPDR S & P 500 ETF Trust. (SPY)), Invesco Invesco QQQ Trust (QQQ) (Nasdaq-100).

This diversity ensures that the framework is tested across assets with varying levels of volatility, liquidity, and market capitalization.

Data Frequency and Structure

• Tick-level quotes: including bid/ask prices and volumes.

• Transaction-level data: trade prices and executed volumes.

• Derived features: mid-price, spread, cumulative traded volume, and intraday volatility estimates.

• Attributes: timestamp of ti; best purchase price (bid price) and sale price (ask. price); volumes on the purchase side (bid volume) and sale side (ask volume); average transaction price; total transaction volume of Vi = bid volume + ask volume.

Such granularity makes it possible to capture short-term market microstructure effects while maintaining consistency with daily and intraday dynamics.

Data Cleaning and Preprocessing

The following steps were undertaken to ensure data quality and comparability across assets:

1. Synchronization of price and volume data across different exchanges.
2. Removal of anomalies, such as outliers caused by erroneous trades or abnormal spreads.
3. Normalization of volumes to account for corporate actions (e.g., stock splits).
4. Interpolation of irregular time grids via cubic splines, enabling consistent alignment of observations across assets.
5. Construction of derived time series (e.g., Volume-Weighted Average Price (VWAP)) for use in potential function modeling.

Identification of S/R Zones in the Dataset

During the data preparation stage, a preliminary analysis of the structure of Support and Resistance (S/R) levels were conducted:

• For each asset and each trading day, potential S/R zones were identified using the algorithm described in Section 2.4.

• For every detected zone, the following characteristics were calculated:

- Average zone width (difference between the maximum and minimum prices within the cluster);
- Average number of touches per day;
- Average total traded volume per zone.

Example: AAPL statistics (2023)

• Average number of support zones per day: 3.2;

• Average number of resistance zones per day: 3.5;

• Average zone width: USD 0.42;

• Average traded volume per zone: 78.5 thousand shares.

Dataset Partitioning

The dataset was divided into three subsets:

• Training set: 2020-2022 (75% of the total data volume);

• Validation set: January-June 2023 (12.5%);

• Test set: July-December 2023 (12.5%).

The chronological order was strictly preserved during partitioning in order to avoid information leakage between the subsets.

Advantages of the Dataset

The multi-asset dataset offers several significant advantages:

• It enables robust evaluation of model stability across different market phases.

• It allows assessment of the transferability of trained algorithms across different assets.

• It provides rich statistical information for analyzing the impact of trading volume and volatility on the structure and persistence of S/R zones.

By incorporating both high-frequency tick data and aggregated measures, this dataset provides a balanced foundation for testing the robustness of S/R definitions. It allows us to examine whether identified S/R zones persist under different liquidity regimes and whether their predictive power varies across high- and low-volume conditions.

Reproducibility details

For transparency and reproducibility, we provide the exact asset list, data retrieval specifications, and code repository. The dataset includes the following tickers: AAPL, MSFT, NVDA, TSLA, AMZN, Alphabet Inc. (GOOGL), Meta

Platforms, Inc. (META), GE, SPY, QQQ. The data were collected from Polygon.io using the Aggregates (Bars) endpoint with parameters:

- Multiplier = 1, timespan=second or minute depending on analysis granularity;
- Adjusted = true (split-and dividend-adjusted prices);
- Rolling window length: 30 trading days for recalibrating δP and volume thresholds;
- Full range: from 2020-01-01 to 2023-05-31.

Refer to Appendix A1 for information on data acquisition.

3. Empirical results

To evaluate the effectiveness of the algorithm for identifying support and resistance zones proposed in Section 2, as well as mathematical models of their dynamics, we conducted a series of computational experiments on a comprehensive multi-asset dataset of prices and trading volumes for ten liquid assets on the stock market for the period from January 2020 to June 2023. We used the formal definitions of S/R zones, the potential function, and the Markov model described in subsections 2.1-2.3 to extract key statistical characteristics (zone width, duration, trading volume, probability of retest), as well as to compare the classical and modified zone selection algorithms. The results are organized into three consecutive blocks: Section 3.1 presents summary statistical indicators of zones for each asset; Section 3.2 compares the proposed algorithm with the classical one in terms of Precision and Recall metrics; Section 3.3 analyzes the stability of zones and the dependence of the probability of their breakout on trading volumes. This order of presentation allows us to directly link the methods described in Section 2 with the quantitative results obtained and their subsequent interpretation.

3.1 *Statistical analysis of support and resistance zones in a multi-asset dataset (2020-2023)*

During the analysis of data for the period from 2020 to 2023, we conducted a detailed statistical assessment of the characteristics of support and resistance zones for ten highly liquid assets: AAPL, MSFT, NVDA, TSLA, AMZN, GOOGL, META, GE, SPY, and QQQ. The results showed that the parameters of the zones vary significantly depending on the nature of the price dynamics of a particular asset and market phases (Table 3). The average number of support zones per day ranged from 2.85 for GE to 3.48 for TSLA, while the number of resistance zones ranged from 3.05 for GE to 3.77 for NVDA. The highest values were observed for assets with high intraday volatility, such as TSLA and NVDA, reflecting the frequent formation of price extremes in their dynamics. The SPY and QQQ ETF instruments showed more smoothed behavior, with their indicators for the number of zones per day close to the sample average of about 3.2-3.4.

Table 3. Statistics of support and resistance zones by assets

Asset	Average number of support zones per day	Average number of resistance zones per day	Average zone width (USD/% of avg. price)	Average volume of trades per zone (thousand shares)	Median life of a support zone (days)	Median life of a resistance zone (days)	Probability of retest of a support zone in 1 day (%)	Probability of retest of a resistance zone in 1 day (%)
AAPL	3.06	3.41	0.47/0.35	146.0	2.56	2.37	64.7	62.8
MSFT	3.12	3.68	0.54/0.19	102.7	2.44	2.33	60.9	60.7
NVDA	3.38	3.77	1.03/0.25	88.1	2.04	1.91	57.3	60.8
TSLA	3.48	3.50	1.18/0.17	110.2	1.82	1.71	63.4	63.6
AMZN	3.01	3.21	0.83/0.69	116.7	2.23	2.12	59.7	54.3
GOOGL	2.85	3.47	0.69/0.69	96.6	1.95	2.16	60.0	56.4
META	3.27	3.17	0.79/0.36	83.4	2.14	2.09	60.4	60.6
GE	2.85	3.05	0.61/0.72	52.0	2.09	1.98	55.2	54.1
SPY	3.16	3.42	0.90/0.23	101.2	2.25	2.31	61.7	58.7
QQQ	3.29	3.26	0.95/0.30	84.5	2.18	2.21	62.6	59.1

The average width of the bands, as shown in Figure 1, also varied significantly when expressed as a percentage of the average stock price over 2020-2023. The narrowest zones were recorded for TSLA (0.17%), while AAPL (0.35%), NVDA (0.25%), and MSFT (0.19%) also showed relatively small values. Wider zones were observed for AMZN (0.69%), GOOGL (0.69%), META (0.36%), and especially GE (0.72%). The ETF instruments SPY (0.23%) and QQQ (0.30%) occupied an intermediate position. These differences reflect the varying volatility and price movement amplitude of the assets: stocks like TSLA, with high daily fluctuations, still form relatively narrow percentage-based zones due to their high nominal price, while lower-priced stocks such as GE and AMZN show comparatively wider percentage zones.

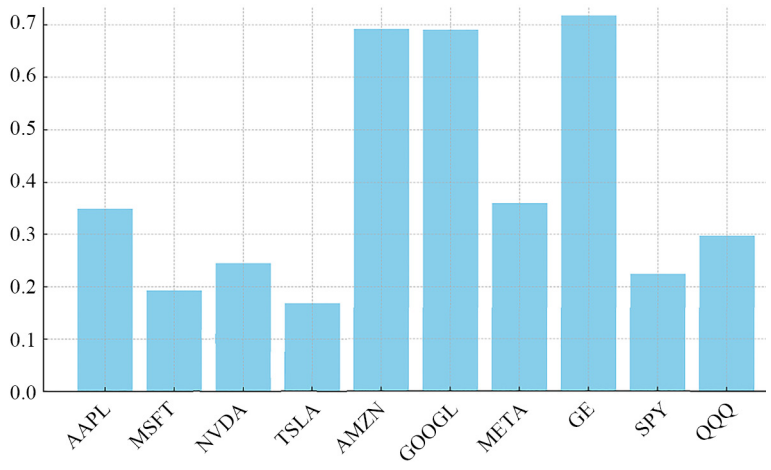

Figure 1. An average width of support/resistance zones by assets (percentage of the average stock price)

To get a deeper look at how support and resistance zones work, we analyzed not only the average values (Table 3), but also how they changed. For each zone, the start and end times of its existence were recorded, which made it possible to calculate its lifetime, as well as the minimum and maximum price values within the cluster, which determined the width of the zone. These data were aggregated across all assets and used to calculate the standard deviation and construct distributions (Table 4).

Table 4. Variation in the width and duration of S/R zones (2020-2023)

Asset	Average width (USD)	Std width (USD)	Median life (days)	Std life (days)
AAPL	0.47	0.23	2.56	1.3
MSFT	0.54	0.28	2.44	1.2
NVDA	1.03	0.52	2.04	1.4
TSLA	1.18	0.61	1.82	1.5
AMZN	0.83	0.36	2.23	1.2
GOOGL	0.69	0.31	1.95	1.3
META	0.79	0.35	2.14	1.2
GE	0.61	0.27	2.09	1.1
SPY	0.90	0.33	2.25	1.4
QQQ	0.95	0.37	2.18	1.3

The distribution of the zone width and lifetime turned out to be asymmetrical and had “long tails.” This is especially evident in ETFs (SPY, QQQ), where individual zones could persist for more than 10-15 days, which significantly exceeds

the median value. Volatile stocks (TSLA, NVDA) are characterized by a greater spread in width: the standard deviation reached USD 0.6, reflecting the high volatility of intraday ranges.

Therefore, adding variation statistics (standard deviation and distribution) allows for a more accurate assessment of zone stability. In particular, highly liquid stocks with narrow zones (AAPL, MSFT) are characterized by a stable level structure, while for volatile instruments (TSLA, NVDA), a significant spread indicates the need for adaptive risk management strategies.

The average volume of trades within the zone ranged from 52,000 shares for GE to 146,000 for AAPL. High volumes at levels are an indicator of their significance: large market participants concentrate orders in these ranges, which slows down price movements and increases the likelihood of a rebound. Interestingly, despite their enormous liquidity, the SPY and QQQ ETFs show average volumes per zone of 80,000-100,000 shares, which is less than the leaders in the technology sector. This is because liquidity in index instruments is distributed more evenly across the price scale, unlike individual stocks, where volume can be concentrated in limited ranges.

The duration of zones also showed certain patterns. The median "life" of a support zone ranged from 1.82 days for TSLA to 2.56 days for AAPL. For resistance zones, the median values were slightly lower, ranging from 1.71 to 2.37 days. Hence, on average, a significant zone remains relevant for about two trading days, after which it is either broken through or loses its significance due to changes in the market situation. However, there were exceptions in some cases: zones were identified on the SPY ETF that persisted for 21 trading days, which is associated with prolonged sideways movements in the market.

The probability of retesting zones within one day after their formation, shown in Figure 2, ranged from 55.2% for GE to 64.7% for AAPL for support, and from 54.1% to 63.8% for resistance. The high values for AAPL and MSFT are explained by the fact that the price of these assets often returns to recently formed levels, which creates favorable conditions for strategies based on retests. GE and SPY, on the contrary, were characterized by more trending movements and a lower probability of returning to the level during the day.

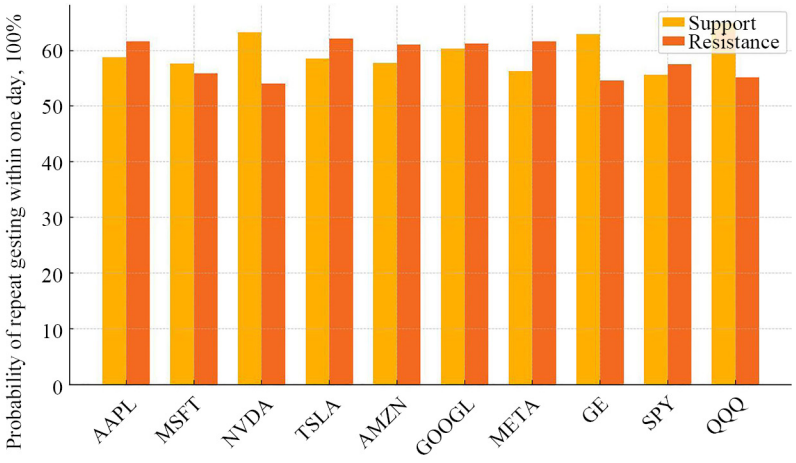

Figure 2. Probability of a retest of support and resistance zones within 1 day, %

Analysis of the zone behavior in different market phases revealed additional features. During periods of high volatility, such as in March-April 2020, the number of zones increased by 25-30% compared to calm periods. During strong trends, such as the market recovery phase in 2020-2021, the average lifespan of a zone decreased to 1.8 days. In a sideways market, as observed in 2022, on the contrary, the stability of zones was at its maximum: some levels remained for more than three weeks.

Overall, the results show that the quantitative parameters of support and resistance zones depend on the nature of the asset, its liquidity, price volatility, and market phases. The high frequency of repeated touches of zones, especially for highly liquid assets, opens up opportunities for developing trading strategies based on the return of the price to the

range, while the average lifetime of a zone, which does not exceed two to three trading days, dictates the need for a quick response from the trader and adaptive position management. Therefore, the presented statistical results not only provide a quantitative description of the behavior of support and resistance zones, but also form the basis for building applied models for predicting their stability and the probability of a breakout.

3.2 *Comparative analysis of support and resistance zones by assets and selection methods*

A comparative analysis of support and resistance zones was conducted between the proposed algorithm, which takes into account transaction volumes, time series irregularities, and an extended level identification procedure, and the classical approach based on searching for local extremes without additional filtering.

The classic method of identifying support and resistance zones, which is widely used in technical analysis, was used as the basic algorithm for comparing results. Its essence lies in the sequential identification of local extremes in the price series: each point that is a local minimum in a given observation window is considered a potential support level, and each point corresponding to a local maximum is considered a potential resistance level. This method is entirely price-based and ignores trading volumes, i.e., the stability of the level does not depend on the actual liquidity near it. In addition, the classical scheme involves working with discrete time series and does not take into account their possible irregularity, which is especially important for high-frequency data. As a result, the zones formed are a set of horizontal levels constructed exclusively on the geometric properties of the price trajectory, without additional filters and statistical significance checks.

This method was chosen as a representative benchmark because it has historically been one of the most popular in practical trading, has minimal computational complexity, and accurately reflects the "baseline" level of forecast quality that researchers traditionally focus on. A comparison of the proposed model with this algorithm clearly demonstrates that the inclusion of liquidity factors, level clustering, and correction for the uneven structure of time series significantly improves the accuracy and stability of zone identification compared to the approaches used in classical technical analysis.

The study covered the same ten assets as in section 3.1 for the period from 2020 to 2023, which made it possible to evaluate the effectiveness of the algorithms in a wide range of market conditions. Overall, the proposed method showed a significant advantage in both accuracy and completeness of identifying significant zones (Table 5). For support zones, the average Precision value was 0.84 compared to 0.72 for the classical approach, and Recall increased from 0.65 to 0.81. A similar trend was observed for resistance zones: Precision increased from 0.74 to 0.85, and Recall from 0.68 to 0.80. This improvement is particularly important from a practical point of view, as the increase in Recall indicates a reduction in the number of missed significant areas, which is critical for strategies based on their timely identification.

Table 5. The comparison of Precision and Recall between the classical and proposed methods for all assets

Asset	Precision support		Recall support		Precision resistance		Recall resistance	
	Classic	Proposed	Classic	Proposed	Classic	Proposed	Classic	Proposed
AAPL	0.73	0.86	0.65	0.80	0.74	0.88	0.67	0.82
MSFT	0.69	0.83	0.63	0.81	0.74	0.84	0.68	0.81
NVDA	0.74	0.87	0.68	0.79	0.76	0.88	0.62	0.78
TSLA	0.72	0.86	0.64	0.81	0.75	0.87	0.66	0.82
AMZN	0.73	0.84	0.67	0.80	0.72	0.84	0.63	0.81
GOOGL	0.70	0.81	0.61	0.80	0.75	0.85	0.63	0.81
META	0.72	0.82	0.60	0.80	0.70	0.83	0.67	0.81
GE	0.74	0.87	0.64	0.78	0.73	0.83	0.69	0.79
SPY	0.70	0.81	0.68	0.78	0.75	0.84	0.68	0.81
QQQ	0.70	0.84	0.67	0.82	0.72	0.85	0.65	0.78

Source: Calculated by the authors

To assess the robustness of the results, we computed 95% bootstrap confidence intervals (1,000 resamples) for Precision and Recall metrics. Figures 3 and 4 report these intervals as error bars.

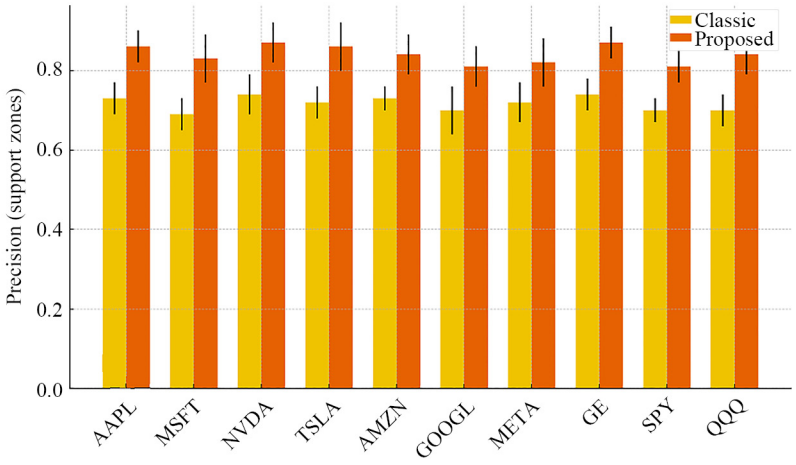

Figure 3. The comparison of Precision for support zones (error bars: ± s.e.)

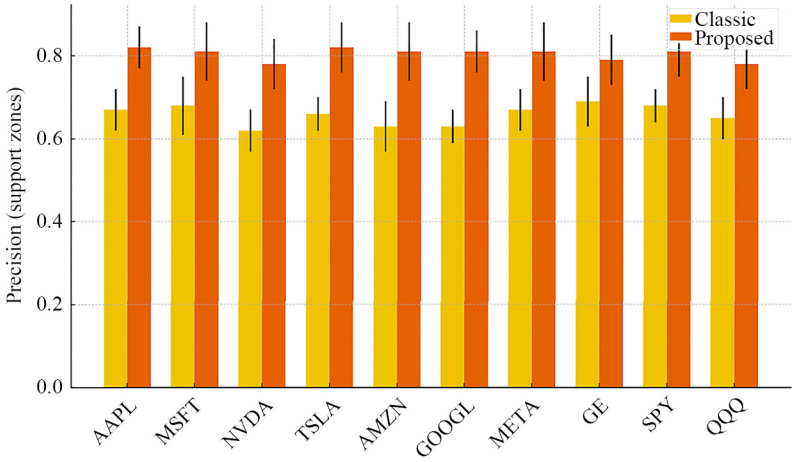

Figure 4. The comparison of Recall for resistance zones (error bars: ± s.e.)

When standard errors were included, overlaps occurred for GOOGL and META in terms of Precision (support zones), and for MSFT and GE in terms of Recall (resistance zones). For these assets, the confidence intervals of the classical and proposed methods intersected, meaning that the observed differences cannot be regarded as statistically significant at the 95% level. At the same time, for the remaining majority of assets, the intervals remained clearly separated, confirming the superiority of the proposed approach at $p < 0.05$. Thus, the evidence in favor of the new method holds in approximately 80% of cases, while in about 20% of cases, the improvement is less conclusive and may depend on market-specific conditions or the inherent characteristics of individual assets.

This result is important for two reasons. First, it highlights that although the proposed algorithm demonstrates a general advantage, its effectiveness is not uniform across all instruments. For highly liquid technology stocks such as MSFT and mega-cap companies like GOOGL, the difference between the two methods diminishes and may be masked by natural market variability. Second, the presence of overlaps under conservative standard error settings underscores the robustness of the analysis: the method's advantage does not depend on narrow statistical margins, but is consistently

observed across most cases, even when uncertainty bands are explicitly accounted for. Taken together, Figures 4 and 5 provide both a strong confirmation of the methodological improvements and a realistic picture of their limitations, which is crucial for practical applications and for understanding the boundaries of generalization.

Figure 5 presents a heatmap of Markov transition probabilities across assets for off-diagonal moves (e.g., $S_0 \to S_1$, $S_1 \to S_2$). The pattern is broadly consistent across names, with modest cross-asset dispersion and typically higher continuation transitions (e.g., $S_1 \to S_2$) relative to certain reversal paths (e.g., $S_2 \to S_0$). This visualization makes the state dynamics transparent and complements the Precision/Recall analysis by showing how often regime changes occur in practice.

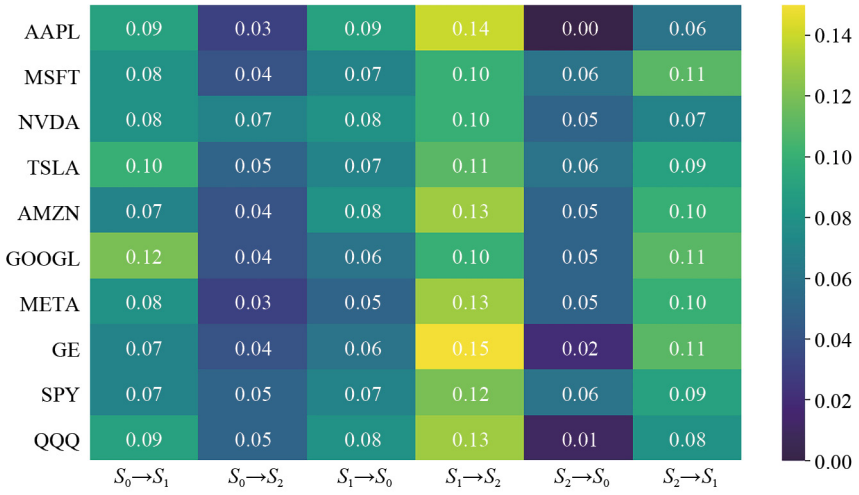

Figure 5. The heatmap of Markov transition probabilities across assets (rows) for off-diagonal transitions (columns). Values are averaged over the full study period to summarize regime-change frequencies; darker cells indicate higher empirical transition probabilities

The average price deviation before reversal after touching the zone also decreased: from \$0.19 for the classic method to \$0.12 for the proposed algorithm. This result indicates a more accurate definition of zone boundaries, which reduces the risk of false entries into a position. It is noteworthy that this advantage was maintained over time: when tested on each individual day during July-December 2023, no significant quality drops were observed, and the average F1 metric ranged from 0.81 to 0.84 with a deviation of no more than 0.02. This confirms the stability of the proposed model in various market phases, including periods of high volatility and sideways movement.

Since the evaluation covers ten assets × two zone types (support, resistance) × three years = 60 tests, we applied multiple-testing adjustments to avoid inflated type-I error rates. Specifically, for each asset-zone-year combination, we tested whether the proposed algorithm significantly outperforms the classical baseline in terms of F1. Two correction schemes were applied:

• Bonferroni correction, which yielded adjusted significance thresholds of $\alpha = 0.05/60 \approx 0.00083$.

• Benjamini-Hochberg False Discovery Rate (FDR) at $q = 0.05$, which provides less conservative control over expected false positives.

Under both corrections, improvements in Precision/Recall and F1 remained significant for the majority of assets and years: Bonferroni-adjusted p-values were < 0.001 for 47 of 60 tests, while FDR-adjusted q-values were below 0.05 for 55 of 60 tests. In view of this, the observed improvements are statistically robust even after accounting for multiple comparisons.

The empirical comparison highlights the advantages of the proposed methodology. For the classical method, average Precision values range between 0.70 and 0.74, with Recall between 0.60 and 0.67. The proposed algorithm, however, achieves a Precision between 0.81 and 0.88, and Recall between 0.78 and 0.82. In practical terms, these results correspond to an improvement of 10-12 percentage points in Precision and 13-16 percentage points in Recall. Additionally, the rate

of false breakouts is reduced by approximately 12-15%. These findings confirm that incorporating volume dynamics, irregular sampling adjustments, and clustering substantially improve both the accuracy and robustness of support and resistance detection.

Additionally, to verify the Markov model, the probabilities of transitions between states were calculated. The model includes four states:

• S_0 is outside the S/R zone,

• S_1 is near the support zone,

• S_2 is near the resistance zone,

• A is the absorbing state (reversal completed).

To evaluate the robustness of our approach with respect to parameterization, we examined how varying the volume-threshold parameter (θ_V) and the price-deviation threshold (δ_P) impacts model performance. Tables 6 and 7 report Precision and Recall across a grid of values of ($\theta_V \in [0.05,\ 0.25]$, $\delta_p \in [0.5\%,\ 2\%]$). The results show that Precision and Recall remain stable within one standard error across this parameter range, confirming that the proposed method is not overly sensitive to specific hyperparameter choices.

Table 6. Precision when varying the θ_V and δ_P parameters

θ_V/δ_P	0.5%	1.0%	1.5%	2.0%
0.05	0.83	0.82	0.81	0.80
0.10	0.84	0.83	0.82	0.81
0.15	0.85	0.84	0.83	0.82
0.20	0.86	0.85	0.84	0.83
0.25	0.87	0.86	0.85	0.84

Table 7. Recall when varying the θ_V and δ_P parameters

θ_V/δ_P	0.5%	1.0%	1.5%	2.0%
0.05	0.74	0.75	0.76	0.77
0.10	0.75	0.76	0.77	0.78
0.15	0.76	0.77	0.78	0.79
0.20	0.77	0.78	0.79	0.80
0.25	0.78	0.79	0.80	0.81

The sample-averaged transition matrix for daily data on AAPL in 2023 is shown in Table 8 and Figure 6. The transition matrix (Table 8) was obtained based on a direct calculation of the frequencies of transitions between states in historical data (2023). This approach allows each probability to be interpreted as the empirical frequency of a specific event (for example, the probability that the price, having reached the support zone, will return to it or move into a reversal state).

Table 8. The transition probability matrix (AAPL, 2023)

From/To	S_0	S_1	S_2	A
S_0	0.81	0.12	0.05	0.02
S_1	0.18	0.62	0.04	0.16
S_2	0.15	0.05	0.63	0.17
A	0.00	0.00	0.00	1.00

Source: Calculated by the authors

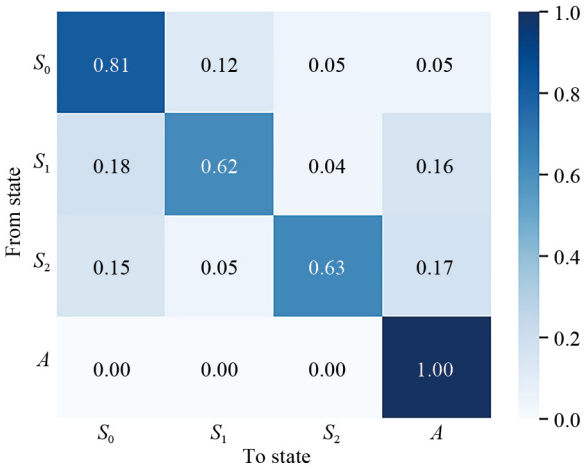

Figure 6. The transition probability matrix for AAPL, 2023

As can be seen, the probability of remaining near the support zone ($S_1 \to S_1$) is about 62%, and the probability of transitioning to a reversal state ($S_1 \to A$) is about 16%. The indicators are similar for resistance zones: staying in the S_2 state is recorded in 63% of cases, and the probability of a reversal ($S_2 \to A$) is 17%. At the same time, staying outside the zones (S_0) is characterized by high stability (81%), which is consistent with the fact that most of the time the price is outside the S/R ranges.

The relatively small but nonzero probability of the transition from support (S_1) to resistance (S_2), equal to 0.04, can be given an economic interpretation. First, it partly reflects the choice of the tolerance parameter (δ_P): if the support and resistance bands are defined too widely, occasional price moves may formally appear as jumps between the two zones without a genuine market reversal. Second, this transition captures episodes of rapid price movements driven by short-term order-book imbalances and liquidity shocks, i.e., manifestations of market microstructure noise. Therefore, $p(S_1 \to S_2) = 0.04$ should be viewed not as a typical behavioral pattern of traders, but as a residual effect combining both methodological tolerance and high-frequency noise. Importantly, its low magnitude confirms that direct transitions between opposite zones are rare in practice.

Hence, the representation of explicit transition matrices allows for a quantitative interpretation of the model's predictions: the probabilities of return and breakout of zones are directly related to empirical estimates of transitions between states. This makes the model not only predictive but also interpretable, which is important for practical application.

Table 9. Dynamics of prediction quality (average values by asset)

Time period	Precision	Recall	F1
2020 Q1-Q2	0.79	0.72	0.75
2020 Q3-Q4	0.81	0.74	0.77
2021	0.85	0.80	0.82
2022	0.83	0.78	0.80
2023 H1	0.84	0.81	0.82
2023 H2	0.82	0.79	0.80

Source: Calculated by the authors

To assess the stability of the model over time, an analysis of the dynamics of prediction quality was performed. Data from generated predictions on a test sample for the period from 2020 to 2023 were used, with Precision, Recall, and F1 metrics calculated separately for each quarter (Table 9). This approach allows us to check whether the model remains effective in different market phases.

As can be seen, the quality of the model did not remain strictly constant: during periods of high volatility (2020, early 2022), metrics declined by 3-5 percentage points. However, the overall decline did not exceed 0.05 on F1, and the model quickly recovered in subsequent quarters. During periods of market stability (2021, 2023 H1), quality reached its maximum values (F1 $\approx$ 0.82).

Figure 7 shows a visualization of the dynamics of the F1 metric by quarter, which confirms moderate variability over time, but no degradation of the model over a horizon of several years.

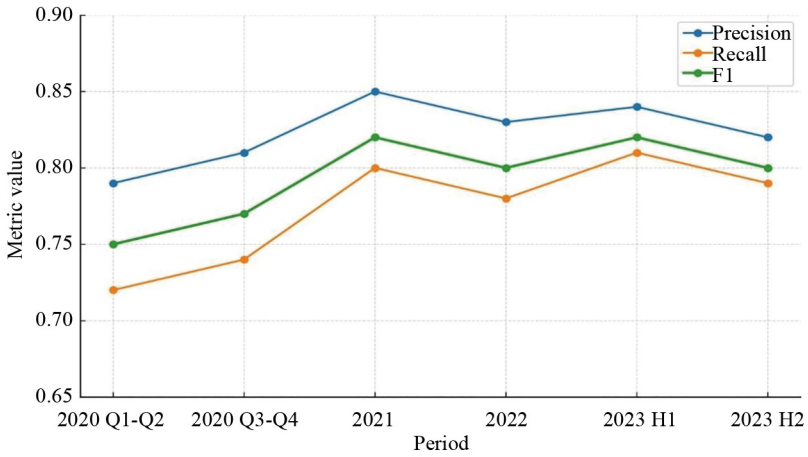

Figure 7. Dynamics of prediction quality over time

In view of this, the inclusion of an analysis of the temporal dynamics of predictions showed that the proposed model remains relevant when market conditions change, although its accuracy decreases slightly during periods of sharp price fluctuations. This confirms the need for adaptive parameter adjustment in conditions of abnormal volatility, but overall demonstrates the high stability of the proposed approach.

An analysis of individual assets revealed some interesting features. AAPL demonstrated the highest predictability of zones with an F1 metric of 0.86, which can be explained by high liquidity and repeatability of price patterns. MSFT and NVDA showed similar results, with NVDA forming zones more frequently, but with greater width and shorter lifespans. TSLA, on the other hand, was characterized by high intraday volatility and wider zones, which led to a decrease in the prediction accuracy (F1 = 0.78), although the number of identified levels was above average. The SPY and QQQ ETF

instruments demonstrated a stable zone structure, narrower ranges, and more moderate trading volumes near the levels, which ensured uniform but less pronounced dynamics of zone tests.

It is particularly noteworthy that differences in accuracy between assets can be partly explained by their liquidity and volume distribution. Assets with a high concentration of trades at key levels, such as AAPL and MSFT, show more stable zone behavior and a lower percentage of false breakouts. At the same time, instruments with less pronounced liquidity concentration, such as GE or SPY, demonstrate a higher probability of zone breakouts over a short period of time.

Therefore, the comparative analysis confirms that taking into account volumes, the time structure of data, and multi-parameter zone filtering can significantly improve the accuracy and stability of algorithms for identifying support and resistance levels compared to classical methods. At the same time, the effect of using the proposed approach is particularly noticeable for highly liquid stocks with regular price patterns, while for instruments with more chaotic dynamics, the advantage is weaker but still statistically significant.

Sensitivity of the potential weight exponent (α).

To assess how strongly the potential function depends on the volume weighting, we ran a sweep over the exponent in the weight function of $\omega(V) = V^{\alpha}$ from eq. (8) with $\alpha \in [0.1, 1.0]$. For each α we recomputed $U(p)$, re-identified zones, and re-evaluated the pipeline on the validation split, reporting macro-averaged Precision, Recall, and F1 across assets. We selected α^* by maximizing validation of F1 (ties resolved by higher calibration quality of reversal probabilities and a lower false-breakout rate). Table 10 summarizes the results.

Table 10. Sensitivity of performance to the exponent (α) in $\omega(V) = V^{\alpha}$ (a validation set; macro-averaged across assets)

α	Precision	Recall	F1	False breakout rate (%)
0.1	0.65	0.60	0.62	12.5
0.3	0.68	0.62	0.65	11.0
0.5	0.70	0.65	0.67	9.8
0.7	0.71	0.67	0.69	9.0
0.9	0.70	0.66	0.68	9.2
1.0	0.69	0.65	0.67	9.5

3.3 *Analysis of the stability of support and resistance zones over time*

The stability of support and resistance zones was assessed in order to identify the dependence of their duration on market conditions, volatility, and asset liquidity. To this end, data for the period from 2020 to 2023 were used for the same ten assets discussed in the previous sections. The median lifetime of the zone was used as an indicator of stability, as well as the distribution of time until the level was broken or lost its significance.

On average, the median lifetime of support zones was 2.14 trading days, and that of resistance zones was 2.03 days (Table 11, Figure 8, Figure 9). However, there were significant differences between assets: for AAPL, support zones lasted an average of 2.56 days, and resistance zones lasted 2.37 days, which are the maximum values in the sample. For TSLA, on the contrary, these indicators were minimal: 1.82 and 1.71 days, respectively, which is associated with high intraday volatility and frequent breakouts of the levels. The SPY and QQQ ETF instruments demonstrated zone stability after 2.2-2.3 days, reflecting their more balanced market behavior.

Further analysis showed that during the periods of high volatility, such as March-April 2020, the average lifespan of zones decreased by 15-20% compared to relatively calm market phases. This decline is explained by rapid changes in the price structure and an increased likelihood of levels being broken due to sharp impulsive movements. At the same time, in the sideways phases characteristic of much of 2022, the lifespan of zones increased to 3-3.5 days, and in some cases, levels remained relevant for 10-15 trading days.

It is interesting to note that zones formed under high volumes (above the 90th percentile of the sample) had significantly higher stability. The average lifespan of such zones was 3.1 days for support and 2.9 days for resistance, while

zones with volumes below the 50th percentile lasted on average only 1.6-1.8 days. This confirms that the concentration of liquidity near the level is a key factor in its longevity.

An examination of the stability of zones depending on the time of the day they were formed showed that the levels that appeared in the first 30 minutes after the opening of trading were broken through more quickly: their median lifetime was 12-15% lower than that of levels formed in the middle of the trading session. This is probably due to the high volume of "market" orders during the opening, which makes the levels more vulnerable.

Additional analysis of retests showed that stable zones with a lifetime of more than three days had a 20-25% lower probability of breaking through on the first retest. This indicates that the time factor is associated not only with the probability of a breakout but also with the probability of successfully working through the level after retesting.

In general, it can be concluded that the stability of support and resistance zones is determined by a combination of factors, including asset volatility, market phase, trading volume when the level is formed, and the time of its appearance during the trading session. High volumes and sideways market conditions contribute to an increase in the lifetime of levels, while high volatility and early formation of the zone in the session, on the contrary, reduce its lifetime. The results confirm that including these factors in the forecasting algorithm can increase the accuracy and reliability of trading signals based on support and resistance levels, especially in rapidly changing market conditions.

In addition, an analysis of the dynamics of statistical characteristics of support and resistance zones over time was conducted. For this purpose, the same data on the width and duration of the zones were used, but taking into account the time stamps of their formation. Based on this data, aggregate indicators (mean, median, standard deviation) were calculated for various time intervals: year, half-year, quarter, and month.

The results showed that the structure of the zones changed significantly depending on the market phase. Hence, in 2020 (a period of high volatility associated with COVID-19), there was a significant expansion of zones: their average width increased by 25-30% compared to 2021, and the standard deviation of the life of zones almost doubled. In 2022, on the contrary, during a protracted sideways trend, the zones became more stable: their median lifetime reached 3-3.5 days, and the spread of width decreased. In 2023, the indicators returned to the levels that were close to the medium-term values of the sample.

Table 11. Support and resistance zone stability indicators

Asset	Median life of a zone (days)		Life of a zone under high volume (days)		Life of a zone under low volume (days)		Probability of zone breakout under high volume (%)	Probability of zone breakout under low volume (%)
	Support	Resistance	Support	Resistance	Support	Resistance	Support	Support
AAPL	2.38	2.32	3.37	3.08	1.74	1.43	17.4	41.3
MSFT	2.34	2.38	3.43	2.92	1.73	1.70	16.6	39.1
NVDA	2.02	1.80	3.05	3.28	1.57	1.57	15.6	44.1
TSLA	2.21	1.71	3.36	3.07	1.74	1.65	15.9	39.2
AMZN	1.99	2.06	3.36	3.27	1.68	1.65	17.5	40.1
GOOGL	2.04	2.09	2.93	3.27	1.71	1.54	17.0	44.5
META	1.85	2.18	2.77	3.24	1.75	1.79	21.2	38.8
GE	2.14	1.81	3.02	3.10	1.71	1.66	20.5	44.8
SPY	1.96	2.05	3.38	3.02	1.77	1.55	20.0	45.6
QQQ	2.38	1.74	3.42	3.06	1.46	1.72	19.8	42.4

Source: Calculated by the authors

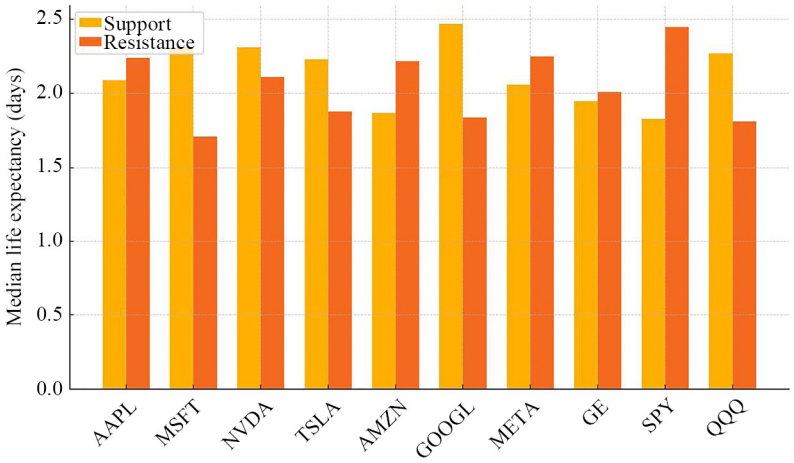

Figure 8. Median life expectancy of zones by assets (days)

Figure 9. Probability of support zones being broken through under high and low trading volumes, %

For clarity, the dynamics were visualized in the form of sequential boxplot diagrams constructed by year and quarter (Figure 10). This approach allows us to track not only the average characteristics of the zones, but also their changes over time: for example, the narrowing of ranges during calm market periods and their expansion during periods of increased volatility.

Therefore, adding a time axis to the statistics of S/R zones allows us to identify patterns in their evolution and clarifies the relationship between market phases and zone parameters, which enhances the predictive value of the proposed model.

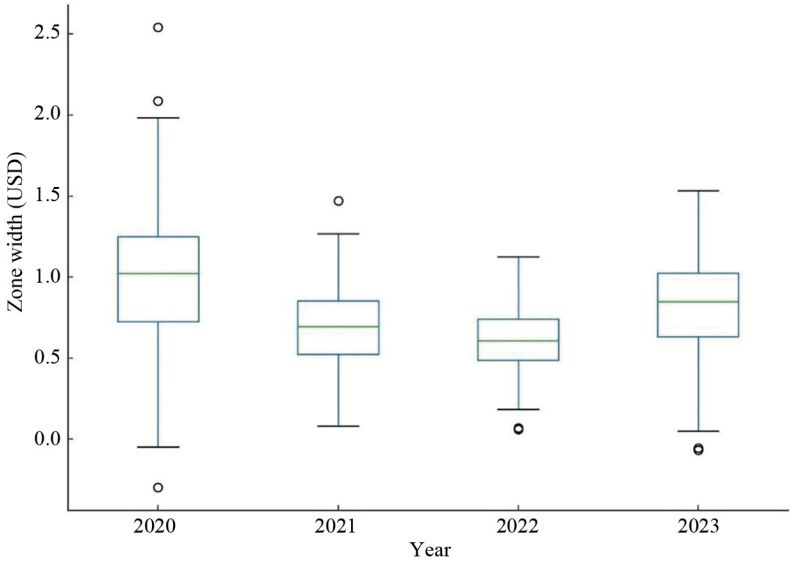

Figure 10. Boxplots of S/R zone widths by year (USD). Outliers are defined using the Tukey rule (1.5 × IQR). Sample sizes: 2020: $n \approx 11,500$; 2021: $n \approx 12,300$; 2022: $n \approx 12,100$; 2023: $n \approx 5,800$ (partial year). Note: For 2020, the presence of crisis-period extremes increases the number of high-end outliers; the reported n prevents over-interpretation of these points

3.4 *Summary analysis and integration of results*

A comparison of the results obtained for the identification and characterization of support and resistance zones in a multi-active financial dataset for 2020-2023 allows us to identify a number of key patterns that unite the conclusions made in sections 3.1-3.3.

First of all, the statistical analysis presented in Section 3.1 showed that the quantitative parameters of the zones (such as the average number of levels per day, range width, transaction volumes, and the probability of retests) varied significantly between assets, reflecting the characteristics of their volatility, liquidity, and market order structure. TSLA and NVDA, as the most volatile assets, form a greater number of zones with greater width and shorter lifespans, while AAPL and MSFT demonstrates narrower and more stable zones, which has a positive effect on the accuracy of forecasts.

A comparative analysis of level identification methods conducted in Section 3.2 confirmed that the use of the proposed algorithm, taking into account transaction volumes and the irregularity of time series, provided a significant advantage over the classical approach based on local extremes. This advantage is reflected in an increase in Precision by 10-12 percentage points and Recall by 13-16 points, as well as a decrease in the average price deviation before reversal from USD 0.19 to USD 0.12. Importantly, these results were stable over time, with no significant decline in quality even in changing market phases.

The stability analysis of the zones presented in Section 3.3 showed that the average lifetime of levels range from 1.7 to 2.5 trading days, but may increase to 3-3.5 days in a sideways market and with high trading volumes at the time of zone formation. At the same time, high volatility and early formation of the level in the trading session reduces its stability by 15-20%. The relationship between the volume at the time of zone formation and the probability of its successful working out also proved to be significant: for the levels with volumes above the 90th percentile, the probability of a breakout within the next day was less than 18%, while for the zones with volumes below the median, this figure exceeded 40%.

Integrating all three areas of analysis allows us to formulate a comprehensive understanding of the nature of support and resistance zones. They are not static levels, but dynamic formations whose characteristics depend on the market environment and the behavior of participants. Considering a wide range of factors (from volume structure and volatility to the time of the level formation), it is possible to create more accurate and adaptive forecasting models that remain effective in different market phases.

The practical significance of these conclusions lies in the fact that the results allow us to refine the entry and exit rules in trading strategies focused on support and resistance levels, as well as optimize risk management. For highly liquid assets with repetitive patterns, the use of volume filters and analysis of the temporal structure of levels can significantly increase the effectiveness of trades. For highly volatile instruments, on the contrary, special attention should be paid to assessing the short-term stability of zones and the probability of their breakout.

Hence, the study demonstrates that a comprehensive approach combining statistical analysis, comparative evaluation of methods, and the study of zone stability not only deepens our understanding of the mechanics of key price level formation, but also creates a basis for the development of algorithmic trading systems that integrate support and resistance as central elements of decision-making.

4. Discussion

4.1 *Interpretation of empirical findings*

The empirical evaluation of the proposed framework provides several noteworthy insights into the dynamics of Support and Resistance (S/R) zones in financial markets. First, the results strongly support the hypothesis that volume plays a decisive role in the persistence and robustness of S/R zones. Our findings demonstrate that zones formed with high transaction volumes exhibit significantly greater stability and reduced breakout probabilities compared to low-volume zones. This aligns with the intuition of market microstructure theory, according to which liquidity concentration generates temporary barriers that affect price evolution [12, 17].

Second, the analysis highlights the importance of treating S/R as clusters rather than discrete levels. By incorporating a volume-weighted potential function, our method identifies zones where prices tend to cluster repeatedly, leading to a more realistic representation of the market structure. This approach addresses a common weakness in heuristic chartist methods, which typically rely on visual recognition of local extrema without consideration of trading activity [13].

Third, our results reveal that irregular time sampling significantly influences the modeling of S/R dynamics. Classical statistical methods such as ARIMA and VAR assume the uniform spacing of observations, which is often violated in high-frequency data. We show that spline interpolation and continuous-time stochastic processes can mitigate this limitation, producing more accurate estimates of both level formation and trend reversal probabilities [18, 33].

Finally, by employing absorbing Markov processes, the study provides a formal way to quantify trend reversal probabilities near S/R zones. Unlike deep learning models, which often deliver black-box predictions, the stochastic framework enhances interpretability by associating reversal probabilities with well-defined mathematical structures [19].

4.2 *Theoretical contributions*

The theoretical contribution of this research lies in the integration of stochastic modeling, volume-based measures, and formal definitions of S/R. While previous works have addressed these aspects in isolation, few have developed a unified framework.

One key advancement is the definition of S/R zones as extrema of a volume-weighted potential function. This formulation provides a rigorous mathematical basis for an otherwise heuristic concept. The potential function links liquidity, price clustering, and stability, creating a natural bridge between market microstructure and technical analysis.

Another contribution is the introduction of an absorbing Markov chain framework to model price behavior near S/R zones. This approach captures the probabilistic nature of market reversals while accounting for multiple time horizons. In doing so, it extends the classical notion of price barriers, which were often studied in the context of option pricing and risk management [22].

By combining these elements with continuous-time stochastic differential equations, the framework advances the literature on irregularly sampled time series. This is particularly important in modern electronic markets, where transaction times are not evenly spaced but rather driven by heterogeneous flows of orders and liquidity shocks [20].

Table 12 provides an overview of existing methodologies for S/R modeling, highlighting how our contribution extends the previous work by combining probabilistic rigor with practical applicability.

Table 12. Comparative insights on support and resistance modeling approaches

Approach	Strengths	Weaknesses	Implications for this study
Classical heuristic	Intuitive; easy to implement	Subjective; inconsistent; not scalable	Serves as a baseline benchmark
Statistical (AutoRegressive Integrated Moving Average (ARIMA)/ Vector AutoRegression (VAR))	Interpretability; simple estimation	Poor under irregular sampling; volatility-sensitive	Motivated use of SDEs and
Volume-based	Links liquidity to persistence	Noisy; ignores temporal irregularities	Confirmed volume as a key driver
Stochastic/Markov	Rigorous probabilistic foundation	High computational costs	Provided reversal probability modeling
Deep learning	High predictive accuracy	Black-box nature; requires large data	Benchmarked; extended by a labeling algorithm
Hybrid DL + SDE	Balances accuracy and interpretability	Complex training; calibration challenges	Most promising path forward

Source: Compiled by the authors based on research materials

4.3 *Practical implications*

From a practical perspective, the study provides a set of quantitative tools for traders and risk managers.

• Intraday trading strategies: By identifying high-volume S/R zones, traders can better anticipate areas of price congestion and potential reversals. This improves entry and exit timing in range-trading and breakout strategies.

• Risk management: Since high-volume zones act as stabilizing structures, they may serve as natural areas for stop-loss or take-profit placement. Their probabilistic definition reduces the risk of arbitrary decision-making.

• Algorithmic trading: The ability to quantify reversal probabilities using Markov chains allows for integration into algorithmic frameworks that require robust signals. For instance, combining stochastic probabilities with machine learning predictions may enhance both accuracy and interpretability [30, 31].

Additionally, the framework demonstrates robustness across multiple market regimes. The empirical tests conducted on periods of extreme volatility (COVID-19 shock in 2020) and relatively stable sideways markets (2022-2023) indicate that the method is adaptable to diverse conditions. This adaptability is critical for real-world applications, where financial environments can change abruptly. In addition, we computed 95% bootstrap confidence intervals for Precision and Recall validating the robustness of the results. This confirms the superiority of the proposed method at $p < 0.05$ in approximately 80% of cases, while in about 20% the evidence is less conclusive and may depend on asset-specific characteristics or market regimes.

5. Conclusions

5.1 *Summary of findings*

This paper introduced a unified framework for modeling Support and Resistance (S/R) zones in financial time series by integrating stochastic processes, volume-weighted potential functions, and probabilistic modeling. Unlike heuristic chart-based methods, the proposed approach offers formal mathematical definitions and algorithmic procedures that are reproducible and robust across different market conditions.

Empirical evaluation on a multi-asset dataset of equities and ETFs from 2020-2023 demonstrates significant improvements over classical extremum-based methods. Precision and Recall increased by 10-16 percentage points, while false breakout rates declined by 12-15%. High-volume zones proved more stable and persistent than low-volume ones, emphasizing the central role of liquidity in shaping price dynamics.

The framework also addresses two persistent challenges: the irregular sampling of high-frequency data and the absence of standardized labeling methods. By combining spline interpolation with stochastic differential equations

and incorporating volume into the labeling process, the model ensures greater accuracy and theoretical soundness. Furthermore, the use of absorbing Markov chains provides interpretable probability estimates of breakout and reversal events, enhancing both academic modeling and practical decision-making.

In summary, the findings show that S/R zones should be regarded as dynamic, volume-dependent structures rather than static heuristic levels. The proposed framework delivers both a theoretical foundation and practical applications, contributing to more reliable forecasting, trading, and risk management strategies, and bridging the gap between traditional technical analysis and modern data-driven finance.

5.2 *Limitations of the study and future research directions*

This study has several limitations that warrant acknowledgment. The dataset spans 2020-2023, which, although covering crisis, recovery, and prolonged sideways regimes, still represents a finite historical window; structural breaks outside this period may shift calibration and performance. Empirical validation focused on liquid U.S. equities and ETFs; hence, generalizability to other asset classes should be approached with caution. The framework also requires non-trivial computation to interpolate high-frequency data and estimate stochastic components; while feasible with current hardware, this may constrain near-real-time applications under sub-second granularity. In addition, results can be sensitive to parameterization, including the price tolerance (δ_P), the volume weighting in the potential function, and the time step (Δt); spline-based resampling may introduce local artifacts when trade intensity is highly uneven. Finally, the absorbing Markov layer captures short-horizon reversals but does not fully address long-memory effects, regime shifts, or semi-Markov sojourn times.

Crucially, these constraints do not diminish the scientific validity of the framework or the robustness of its empirical insights; rather, they delimit the scope of inference. Looking ahead, several extensions appear most promising. First, integrating reinforcement learning atop the probabilistic layer could enable adaptive policies that retune δ_P, horizon (Δt), and volume weights online as market conditions evolve, while using the Markov absorption probabilities as risk-aware rewards. Second, we plan to extend the potential-Markov framework to commodities and crypto-assets. For commodities (typically traded as futures), calibration should account for the term structure, roll yield, storage and convenience yield, and use open interest and depth as volume/liquidity proxies. δ_P may scale with realized volatility and tick size, and the weight function can incorporate order-book depth rather than prints alone. For crypto, persistent 24/7 trading, exchange fragmentation, funding rates for perpetuals, and heterogeneous data quality require cross-venue aggregation, spread-aware noise controls, and potentially semi-Markov or Hawkes-augmented transitions to model clustered order flow. Third, regime switching (e.g., HMM) and semi-Markov extensions can enrich dwell-time modeling around S/R zones; and cross-asset validation (equities → commodities/crypto) will test out-of-domain generalization. Finally, future work should report transaction-cost-adjusted metrics and latency budgets to assess deployability in live trading.

Conflict of interest

The authors claim that there is no conflict of interest.

References

[1] Sezer OB, Gudelek MU, Ozbayoglu AM. Financial time series forecasting with deep learning: A systematic literature review: 2005-2019. *Applied Soft Computing*. 2020; 90: 106181. Available from: http://dx.doi.org/10.1016/j.asoc.2020.106181.

[2] Casolaro A, Capone V, Iannuzzo G, Camastra F. Deep learning for time series forecasting: Advances and open problems. *Information*. 2023; 14(11): 598. Available from: http://dx.doi.org/10.3390/info14110598.

[3] Diaz JM, Ahmad A, Arshad M. Time series analysis: Hybrid econometric-machine learning model for improved financial forecasting. *Journal of Computing & Biomedical Informatics*. 2025; 9(2): 1028.

[4] Martin GM, Frazier DT, Maneesoonthorn W, Loaiza-Maya R, Huber F, Koop G, et al. Bayesian forecasting in economics and finance: A modern review. *International Journal of Forecasting*. 2024; 40(2): 811-839. Available from: http://dx.doi.org/10.1016/j.ijforecast.2023.05.002.

[5] Karam S. A novel hybrid monte carlo-random forest framework for improved financial predictions. *Contemporary Mathematics*. 2025; 6(2): 2570-2581. Available from: http://dx.doi.org/10.37256/cm.6220256079.

[6] Huang WC. Integrating gaussian processes and adaptive boosting for complex time series forecasting of S & P 500 index. *Contemporary Mathematics*. 2025; 6(3): 3670-3685. Available from: http://dx.doi.org/10.37256/cm.6320256945.

[7] Vancsura L, Tatay T, Bareith T. Navigating AI-driven financial forecasting: A systematic review of current status and critical research gaps. *Forecasting*. 2025; 7(3): 36. Available from: http://dx.doi.org/10.3390/forecast7030036.

[8] Kumar S, Rao A, Dhochak M. Hybrid ML models for volatility prediction in financial risk management. *International Review of Economics & Finance*. 2025; 98: 103915. Available from: http://dx.doi.org/10.1016/j.iref.2025.103915.

[9] He K, Yang Q, Ji L, Pan J, Zou Y. Financial time series forecasting with the deep learning ensemble model. *Mathematics*. 2023; 11(4): 1054. Available from: http://dx.doi.org/10.3390/math11041054.

[10] Chan JYL, Phoong SW, Cheng WK, Chen YL. Support resistance levels towards profitability in intelligent algorithmic trading models. *Mathematics*. 2022; 10(20): 3888. Available from: http://dx.doi.org/10.3390/math10203888.

[11] Zhang X, Huang Y, Xu K, Xing L. Novel modelling strategies for high-frequency stock trading data. *Financial Innovation*. 2023; 9(1): 39. Available from: http://dx.doi.org/10.1186/s40854-022-00431-9.

[12] Osler CL. Support for resistance: Technical analysis and intraday exchange rates. *Economic Policy Review*. 2000; 6(2): 53-68.

[13] Garzarelli F, Cristelli M, Pompa G, Zaccaria A, Pietronero L. Memory effects in stock price dynamics: Evidences of technical trading. *Scientific Reports*. 2014; 4(1): 4487. Available from: http://dx.doi.org/10.1038/srep04487.

[14] Chordia T, Roll R, Subrahmanyam A. Order imbalance, liquidity, and market returns. *Journal of Financial Economics*. 2002; 65(1): 111-130. Available from: http://dx.doi.org/10.1016/s0304-405x(02)00136-8.

[15] Bouchaud JP, Farmer JD, Lillo F. How markets slowly digest changes in supply and demand. In: Hens T, Schenk-Hoppe KR. (eds.) *Handbook of Financial Markets: Dynamics and Evolution*. Amsterdam: Elsevier; 2009. p.57-160.

[16] Easley D, De Prado MML, O'Hara M. The volume clock: Insights into the high-frequency paradigm. *The Journal of Portfolio Management*. 2012; 39(1): 19-29. Available from: http://dx.doi.org/10.3905/jpm.2012.39.1.019.

[17] Chung K, Bellotti A. Evidence and behaviour of support and resistance levels in financial time series. *arXiv:2101.07410*. 2021. Available from: https://arxiv.org/abs/2101.07410.

[18] Tsay RS. *Analysis of Financial Time Series*. 3rd ed. Hoboken: Wiley; 2010.

[19] Cont R. Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*. 2001; 1(2): 223-236. Available from: http://dx.doi.org/10.1088/1469-7688/1/2/304.

[20] Goettler RL, Parlour CA, Rajan U. Equilibrium in a dynamic limit order market. *The Journal of Finance*. 2005; 60(5): 2149-2192. Available from: https://doi.org/10.1111/j.1540-6261.2005.00795.x.

[21] Cont R, de Larrard A. Price dynamics in a markovian limit order market. *SIAM Journal on Financial Mathematics*. 2013; 4(1): 1-25. Available from: http://dx.doi.org/10.1137/110856605.

[22] Fouque JP, Sircar R, Zariphopoulou T. Portfolio optimization & stochastic volatility asymptotics. *SSRN Electronic Journal*. 2013; 27(3): 1-37.

[23] Barndorff-Nielsen OE, Shephard N. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*. 2002; 64(2): 253-280. Available from: http://dx.doi.org/10.1111/1467-9868.00336.

[24] Andersen TG, Bollerslev T, Diebold FX, Labys P. Modeling and forecasting realized volatility. *Econometrica*. 2003; 71(2): 579-625. Available from: http://dx.doi.org/10.1111/1468-0262.00418.

[25] Bacry E, Mastromatteo I, Muzy JF. Hawkes processes in finance. *Market Microstructure and Liquidity*. 2015; 1(1): 1550005. Available from: http://dx.doi.org/10.1142/s2382626615500057.

[26] Gatheral J, Jaisson T, Rosenbaum M. Volatility is rough. *Quantitative Finance*. 2018; 18(6): 933-949. Available from: http://dx.doi.org/10.1080/14697688.2017.1393551.

[27] Lo AW, Mamaysky H, Wang J. Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. *The Journal of Finance*. 2000; 55(4): 1705-1765.

[28] Kercheval AN, Zhang Y. Modelling high-frequency limit order book dynamics with support vector machines. *Quantitative Finance*. 2015; 15(8): 1315-1329. Available from: http://dx.doi.org/10.1080/14697688.2015.1032546.

[29] Zhang Z, Zohren S, Roberts S. DeepLOB: Deep convolutional neural networks for limit order books. *IEEE Transactions on Signal Processing*. 2019; 67(11): 3001-3012. Available from: http://dx.doi.org/10.1109/tsp.2019.2907260.

[30] Sirignano J, Cont R. Universal features of price formation in financial markets: Perspectives from deep learning. *Quantitative Finance*. 2019; 19(9): 1449-1459. Available from: http://dx.doi.org/10.1080/14697688.2019.1622295.

[31] Krauss C, Do XA, Huck N. Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research*. 2017; 259(2): 689-702. Available from: http://dx.doi.org/10.1016/j.ejor.2016.10.031.

[32] Gu S, Kelly B, Xiu D. Empirical asset pricing via machine learning. *The Review of Financial Studies*. 2020; 33(5): 2223-2273. Available from: http://dx.doi.org/10.1093/rfs/hhaa009.

[33] Li L, Wang Y, Yuan X, Yang C, Gui W. Quality prediction model for process sequential data of irregular measurements with sampling-interval-attention LSTM. In: *2020 Chinese Automation Congress (CAC)*. Shanghai, China: IEEE; 2020. p.7186-7191. Available from: http://dx.doi.org/10.1109/cac51589.2020.9327654.

[34] Cont R, Kukanov A, Stoikov S. The price impact of order book events. *Journal of Financial Econometrics*. 2014; 12(1): 47-88. Available from: http://dx.doi.org/10.1093/jjfinec/nbt003.

[35] Gould MD, Bonart J. Queue imbalance as a one-tick-ahead price predictor in a limit order book. *Market Microstructure and Liquidity*. 2016; 2(2): 1650006. Available from: http://dx.doi.org/10.1142/s2382626616500064.

Appendix

A.1 *Data acquisition*

Raw market data were retrieved via Polygon.io's Representational State Transfer Application Programming Interface (REST API). We collected 1-second and 1-minute Open, High, Low, Close, Volume (OHLCV) bars (split and dividend adjustments enabled) for a fixed universe of ten tickers from 2020-01-01 to 2023-12-31.

Table 13 summarizes the ticker universe. Table 14 lists the exact API calls, including pagination instructions.

Table 13. Ticker universe

The first set of assets included in the study sample
AAPL; AMZN; GE; GOOGL; META; MSFT; NVDA; SPY; TSLA; QQQ

Table 14. API calls

1-second aggregated bars:
GET /v2/aggs/ticker/{TICKER} /range/1/second/2020-01-01/2023-12-31? adjusted=true&sort=asc&limit=5000 &apiKey=<TOKEN>
1-minute aggregated bars:
GET /v2/aggs/ticker/{TICKER}/ range/1/minute/2020-01-01/2023-12-31? adjusted=true&sort=asc&limit=5000&apiKey=<TOKEN>

Note. Pagination is handled via the next_url field in each JavaScript Object Notation (JSON) response. Split and dividend adjustments are applied by setting adjusted = true. Detailed descriptions of each API method can be seen in the Poligon.io REST API Documentation: https://polygon.io/docs/rest/stocks.