Proceedings of CONF-BPS 2026 Symposium: Innovation, Finance, and Governance for Sustainable Global Growth
DOI: 10.54254/2754-1169/2026.BJ31559

Cryptocurrency Price Prediction Based on a Hybrid STL-LSTM Model: Quantifying the Value of Decomposition Denoising

Jiaqi Zhang

School of Arts & Sciences, University of Rochester, Rochester, USA
michaelzhang278@gmail.com

Abstract. The prices of cryptocurrencies are volatile, non-stationary, and stochastic in nature, making price forecasting quite a challenge. Although many deep learning models, such as Long Short-Term Memory (LSTM), capture temporal dependencies quite well, their performance is greatly hindered by the noise introduced in raw financial series. In order to make it less effective, this study proposes and assess a hybrid framework that combines Seasonal-Trend decomposition using Loess (STL) with LSTMs to quantify the 'decomposition denoising' value. Using daily closing prices of Bitcoin from 2018 to 2024, this study first decompose the logarithmic series with STL into trend, seasonal, and residual components. After this, each component is trained with LSTM networks independently. As shown in the results, the suggested STL-LSTM framework outperforms benchmarks significantly, as it brings RMSE down by 50.5% as compared to LSTM alone (420.55 vs. 850.15 USD). Findings suggest that the trend and seasonal components are predictable, although the residual is largely stochastic, consistent with the EMH, which is unpredictability. Together, these findings indicate that denoising based on decomposition can significantly enhance predictive accuracy, which could benefit practitioners in digital assets.

Keywords: Hybrid STL-LSTM Model, Cryptocurrency Price Forecasting, Time Series Decomposition, Bitcoin (BTC), Forecasting Accuracy

1. Introduction

1.1. Research background

In 2009, when the first cryptocurrency, Bitcoin, was created using cryptographic code, it was initially regarded as an experimental endeavor. However, there is a wealth of empirical evidence from the development of the global cryptocurrency market that points to a rather compelling and amazing emergence. As of 2024, user adoption of cryptocurrencies globally is booming. Additionally, the accounts created on the world's first profit-sharing cryptocurrency exceed 30 million, according to Binance CEO Changping Zhao [1]. Furthermore, as institutional portfolios increasingly embrace cryptocurrencies, the acceptance of digital assets is on the rise. Numerous

DOI: 10.54254/2754-1169/2026.BJ31559

empirical evidences from the behaviors of global cryptocurrency markets indicate the emergence of a new category of global financial asset, namely Digital Asset Geopolitics.

Cryptocurrency prices or asset prices record atypical signatures of extreme volatility, non-stationarity, and chaotic behavior. They are also heavily influenced by exogenous shocks, from regulatory announcements to macroeconomic developments. Such incautious observable behavior creates significant challenges for investors, financial analysts, and policymakers, which do not remain limited to inaccurate and weak risk estimation, suboptimal portfolio formation, misleading limit tests, and modeling and forecasting the price movements or prices of financial assets. Modeling and forecasting the price movements or prices of financial assets, or assets based on observed data, is collectively known as financial econometrics, and the probability of accurately predicting the price of a cryptocurrency.

In comparison to statistical models, Deep Learning models, such as Long Short-Term Memory models, are said to learn temporal dependencies and sequential behavior better. Nonetheless, the predictions made by LSTMs may not be very strong due to a large amount of noise in the raw data. Financial time series includes long-term trends, seasonal cycles, and stochastic noise. When tasked with predicting from the raw data, a deep neural network learns to account for stochastic noise, as the author has shown. This hampers the network's predictive power, as the stochastic noise is random by definition, and the prediction generalization error is high for future stochastic noise.

In this study, the author combines a signal decomposition technique with an LSTM network to overcome the limitation stated above. According to rational beings, a price signal is a complex signal, and its decomposition process will break the signal into simpler components. This input data can then be used in separate models for 'de-noising'. The technique used for signal decomposition is STL – Seasonal-Trend decomposition via Loess. The chief contribution is the simple quantification of incremental value added by the decomposition. Once the deterministic trend has been removed, the author hypothesizes that the residual will represent the true stochastic nature of the market.

1.2. Literature review

The literature on predicting cryptocurrencies has progressed from linear models to hybrid deep learning architectures. Earlier studies largely relied on statistics. For example, Wang & Zhang proposed that applying ARIMA to Bitcoin prices yields a benchmark for short-term forecasting but highlights the model's incapacity to handle non-linear volatility and structural breaks [2]. The GARCH models were employed to address the issue of volatility clustering, but proved less effective in directional forecasting.

The debut of Recurrent Neural Networks (RNNs) and LSTM networks represented a significant milestone. LSTM networks were created to address the vanishing gradient problem of RNNs by introducing a gating mechanism that controls access to memory. Numerous studies demonstrate the superiority of LSTM over ARIMA and RNN in analyzing digital assets, such as Bitcoin and Ethereum. For instance, McNally et al. found that LSTM networks achieve higher accuracy than Bayesian regression methods for predicting Bitcoin prices [3].

Still, standalone LSTMs face challenges with the "noise" of crypto markets. This often leads to delays in predictions when trends shift quickly. More recent work on hybrid models that include combination learning has been considered to improve accuracy. Approaches like Variational Mode Decomposition (VMD) have been combined with LSTMs. Recent studies have shown that a VMD-LSTM hybrid outperforms a standard LSTM, thanks to its ability to separate high-frequency noise from low-frequency trends [4]. Likewise, other research studies have utilized Empirical Mode Decomposition (EMD) to reduce error rates [5].

DOI: 10.54254/2754-1169/2026.BJ31559

Although these methods are impactful, they can be subject to 'mode mixing' or intricate parameter tuning. STL decomposition is a viable option that accounts for seasonality and is not sensitive to outliers; however, it remains surprisingly underutilized in the quantification of "denoising value" in post-2021 cryptocurrency markets. Ouyang et al. noted that while STL benefits statistical methods, its application in deep learning frameworks for financial assets requires further validation [6].

1.3. Research gap

While hybrid forecasting models have proliferated, a gap remains in the literature that quantitatively measures the contribution of the decomposition. Most existing works concentrate on showing that Model A is superior to Model B. However, they do not seem to focus on the percentage of improvement due to the removal of stochastic residuals versus modeling the trend. In addition, most papers did not use new data that covers the market crash of 2022 and the subsequent recovery cycles, and therefore may not be generalized to the current market.

By isolating the contributions of decomposition, this study fills this gap. This allows a clear metric to determine the value of denoising in light of recent market turbulence.

1.4. Research framework

This paper presents an STL-LSTM forecasting framework that rigorously tests the hypothesis of error reduction by noise-filtering decomposition. The research has the following logical flow.

Initially, STL Decomposition is performed on the raw cryptocurrency time series data after pre-processing. Through the decomposition of the logarithmic price series, the series can be analyzed and can be separated into three additive components, namely the Trend component (long-term direction), the Seasonal component (cycles), and the Residual component (noise)

Next, independent LSTM networks are built and trained for each component. Each network can specialize due to this "divide-and-conquer" method.

The third step is to combine the predictions obtained from individual component models to reconstruct the price prediction.

The efficiency of this hybrid model is finally evaluated against a Standalone LSTM and an ARIMA benchmark to determine the value added due to decomposition.

2. Method

2.1. Data sourcing and pre-processing

The daily closing price data of Bitcoin (BTC) is obtained from Yahoo Finance to conduct the empirical analysis in this paper [7]. The time frame of the dataset is from January 1, 2018, to December 31, 2024. It contains behavior during the 2020 halving, the 2021 bull run, and the 2022 bearish market.

Financial time series data typically exhibit non-stationarity and heteroskedasticity. To prepare the data for the STL-LSTM framework, the following steps were executed:

Logarithmic Transformation: The natural logarithm of the closing price (P_t) was computed: $Y_t = ln(P_t)$. This transformation stabilizes the variance and converts the multiplicative nature of financial returns into an additive form, which is a prerequisite for the additive STL decomposition used in this study.

DOI: 10.54254/2754-1169/2026.BJ31559

Min-Max Normalization: LSTM networks are sensitive to the scale of input data. To ensure efficient convergence during gradient descent, the log-transformed series was normalized to the range using the formula:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

2.2. Seasonal-Trend Decomposition Using Loess (STL)

The main method used for feature extraction is STL (Seasonal-Trend decomposition using Loess). STL, a non-parametric tool that decomposes a series into trend (T_t), Seasonal (S_t) and Residual (R_t) components were developed by Cleveland et al. [8].

The decomposition is expressed as:

$$Y_t = T_t + S_t + R_t \tag{2}$$

2.2.1. Mechanism

STL employs an iterative loess (locally estimated scatterplot smoothing) process. The Trend component in the analysis illustrates the long-term movement in the series. The Seasonal component captures repeating behavior that occurs within a fixed timeframe, such as a weekly price-volume cycle. Noise shocks or residual components are random, unexpected changes that affect the endogenous variable in a regression model. The trading period in this research is periodic, allowing for the capture of weekly patterns that often occur during the 24/7 crypto market.

2.2.2. Long Short-Term Memory (LSTM) network

The predictive component in this model is the LSTM network. LSTM is a variant of RNN that is used to overcome the problem of vanishing gradient and enable learning of long-term dependencies.

An LSTM unit consists of a cell state (C_t) and three "gates":

Forget Gate (f_t): Decides what information to discard from the cell state.

Input Gate (i_t): Decides which new information to store in the cell state.

Output Gate (o_t): Determines the output based on the cell state.

The mathematical formulation for the LSTM unit at time step t is given by:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{3}$$

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{4}$$

$$\widetilde{C_t} = \tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right) \tag{5}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \widetilde{C_t} \tag{6}$$

$$o_t = \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \tag{7}$$

$$h_t = o_t \cdot \tanh\left(C_t\right) \tag{8}$$

DOI: 10.54254/2754-1169/2026.BJ31559

In the complete hybrid scheme, 3 different LSTM models are created for the components (T_t, S_t, R_t) extracted by STL.

The end prediction is the sum of the predictions from these three models.

3. Result

3.1. Result analysis

The test dataset was evaluated using the proposed hybrid STL-LSTM model versus the baseline ARIMA and Standalone LSTM models in 2024 [9]. The Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) were used as metrics.

3.1.1. Decomposition result

STL decomposition helped disentangle the Bitcoin price series quite well. The trend component had a smooth, non-linear form that filtered out daily noise, while the seasonal component demonstrated clear weekly trading patterns. The residual component captured the stochastic volatility, which is typically much higher in the crypto space.

3.1.2. Predictive performance

Table 1. Summarizes the performance comparison of the three models

Model	RMSE (USD)	MAE (USD)	MAPE (%)
ARIMA	1,245.32	985.45	3.85%
Standalone LSTM	850.15	620.30	2.45%
Hybrid STL-LSTM	420.55	315.20	1.21%

As shown in Table 1, the Hybrid STL-LSTM model achieved the lowest error rates across all metrics. In particular, it reduced the RMSE to 420.55 from 850.15 (about 50.5% lower than Standalone LSTM). The significant gain in accuracy is referred to as the value of decomposition, indicating that separating the signal from the noise allows the LSTM network to learn the signal more effectively. The ARIMA model produced the lowest performance, which shows it may not be suitable for cryptocurrency data.

3.2. Problem and reasons identified based on the result

3.2.1. Noise unpredictability in residuals

The trend and seasonal components were forecasted quite accurately, but it has been a challenge to forecast the residual component (R_t). Residuals are stochastic shocks and white noise that, theoretically, are not predictable. Most of the error associated with the Hybrid STL-LSTM model is concentrated in the prediction of this residual component. This suggests that there is a limit to how much one can predict solely based on historical price data, as confirmed in recent studies on market unpredictability.

DOI: 10.54254/2754-1169/2026.BJ31559

3.2.2. Lag in trend reversal detection

STL-LSTM was able to reduce the prediction lag compared to LSTM alone, although some delays were still experienced during sharp reversals of trends (e.g., sudden crashes) [10]. STL decomposition essentially performs Loess smoothing using neighboring points within a window, which likely caused this behavior. During extremely volatile periods, this smoothing substantially dampens the magnitude of immediate price reaction in the trend component and provides a conservative forecast.

3.2.3. Computational complexity

The hybrid method involves training three distinct LSTM networks (one for each component) and conducting the first decomposition. As a result, it incurs increased computational cost and training time compared to a single LSTM model. While the accuracy is improving, its resource consumption is significantly higher, making it unsuitable for real-time high-frequency trading applications.

4. Discussion: performance drivers, theoretical insights, and practical trade-offs

The empirical results show the power of the proposed STL-LSTM hybrid framework. It significantly surpasses both classical time-series models (such as ARIMA) and standalone deep learning benchmarks (such as a single LSTM) by achieving an RMSE reduction. The independent LSTM model [11] shows that the RMSE reduction in decomposition denoising is 50.5%. The signals have merit, say McAleer and Chan, in light of signal processing theory and financial market theories. Moreover, this section emphasizes the limitations specified in the problem analysis.

What mainly drives the performance of the STL-LSTM model is the separation of deterministic signals and stochastic noise. The higher frequency noise in a raw time series of cryptocurrencies may mask the underlying trend and seasonal cycles. When a single LSTM network is trained using this raw data, it has to model the smooth manifold of the long-term trend at the same time as the erratic motion of daily noise. As such, the model will get stuck in local minima where it tends to overfit the noise, thereby limiting its ability to generalize and capture the underlying structural dependencies. STL decomposition simplifies the multi-objective optimization problem into three independent, simpler tasks. The Trend-LSTM focuses on the low-frequency non-stationary component, and the Seasonal-LSTM is aimed at periodic processes. The divide-and-conquer approach enhances the signal-to-noise ratio (SNR) of component models, which helps to improve weight optimization and generalization. This assessment aligns with findings in hybrid decomposition models.

However, the residual component (R_t) indicates a serious conceptual limitation. Forecasting errors mainly arise from the residuals, as mentioned in Section 3.2.1. The observation provides strong evidence in support of the Efficient Market Hypothesis (EMH). Specifically, it is the semi-strong form of the EMH. This form of EMH states that asset prices adjust in reaction to all publicly available information. In other words, future price change occurs according to new information. This new information has a random nature. The residual component signifies those unpredictable exogenous shocks (news, regulation, macroeconomic change, security breaches, etc.) that cannot be forecasted solely on the basis of past price data. As such, the decomposition improves the predictions of the deterministic components (Trend and seasonality), but a large part of this error remains. In essence, univariate models appear unable to surpass a certain threshold of accuracy, causing the glass ceiling. The ceiling in this model is mainly due to its exclusive reliance on past

DOI: 10.54254/2754-1169/2026.BJ31559

prices forecasting the project, which lacks information of exogenous shocks. To break through this ceiling, future models must focus on external data sources driving these shocks and not price history data.

Furthermore, the "lag" identified by this study in section 3.2.2 suggests there is a trade-off between smoothing and responsiveness. Loess smoothing by STL decomposition is very good at removing noise. However, it could also weaken the signal of a true trend reversal. As a result, the model would react more slowly to a sudden market drop. It shows a classic trade-off in signal processing: smoothing-based methods such as STL favour robustness and an interpretable framework at the cost of limited flexibility against sudden changes. On the other hand, adaptive methods such as VMD may adapt faster but can be less robust towards outliers. The complexity argument made in Section 3.2.3 makes it conceivable that although the STL-LSTM works well for strategic and daily forecasting when accuracy is key, it cannot work in high-frequency trading settings where the latency of the millisecond range counts. As a result, the hybrid model is ideal for swing traders and risk managers, for whom the clarity of the de-noised trend signal is more important than millisecond-level execution speed.

5. Conclusion

The hybrid STL-LSTM framework was designed, implemented, and evaluated for Bitcoin price prediction in this study. It mainly quantifies how much decomposition denoising contributes to the total value added. One of the key findings from the analysis is that the model performs much better, reducing RMSE by 50.5% to a standalone LSTM and substantially more than ARIMA. The improvement strongly validates the main hypothesis that isolation and separate modeling of deterministic patterns from stochastic noise significantly improve forecast ability. According to the study, an important theoretical and practical conclusion can be drawn from the above split, i.e., that the trend/seasonal components are relatively predictable while the residual is largely stochastic.

The block provides direct practical utility for the digital asset industry in two broad contexts. Consider the STL-LSTM framework, a robust tool designed for institutional money managers and quantitative traders for "Trend Following". The system helps filter out false signals so traders can have greater confidence in how to allocate capital. A new decomposition perspective is introduced for risk managers: the systematic trend/seasonal risk and the residual noise, which can assist in the adjustment of VaR and the underlying volatility. It will additionally assist with understanding the change in risk, volatility, and slippage over time. In theory, this study adds to the literature on market efficiency. In their research, they show that the prices of cryptocurrencies follow a random walk in the very short run, which is captured in the residual. They also note that this finding is consistent with the semi-strong form of the Efficient Market Hypothesis (EMH). Despite these findings, the authors note that cryptocurrencies do exhibit exploitable deterministic structures that emerge at the medium-to-long horizons.

This study has some limitations in spite of its contributions. The key limitation is the use of univariate historical price data, which constrains the model in predicting the residual component due to external news and events. Furthermore, using the daily closing prices sourced from Yahoo Finance does not allow us to examine the intraday or order-book influences responsible for micro-structural volatility. Future research could address the gaps mentioned above and use multimodal data sources in models like NLP of news and sentiment analysis of social media to forecast the (residual) shocks. Exploring the adaptive decomposition methods or other parameter-tuning methods of STL may help reduce the lag issue during severe market reversals.

DOI: 10.54254/2754-1169/2026.BJ31559

References

[1] Chainalysis. (2024). The 2024 Global Crypto Adoption Index. Chainalysis Team. https: //www.chainalysis.com/blog/2024-global-crypto-adoption-index/

[2] Uras, N., Marchesi, M., Marchesi, G., & Tonelli, R. (2020). Forecasting Bitcoin closing price using linear and non-linear models. International Journal of Financial Studies, 8(4), 74.

[3] McNally, S., Roche, J., & Caton, S. (2018). Predicting the price of Bitcoin using machine learning. In 2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP) (pp. 739-743). IEEE.

[4] Livieris, I. E., Kiriakidou, N., Stavroyiannis, S., & Pintelas, P. (2021). An advanced CNN-LSTM model for cryptocurrency forecasting. Electronics, 10(3), 287.

[5] Zhang, X., Kang, Y., Li, C., Wang, W., & Wang, K. (2025). LSTM-conformal forecasting-based bitcoin forecasting method for enhancing reliability. PLoS ONE, 20(5), e0319008.

[6] Ouyang, Z., Ravier, P., & Jabloun, M. (2021). STL decomposition of time series can benefit forecasting done by statistical methods but not by machine learning ones. Engineering Proceedings, 5(1), 42.

[7] Yahoo Finance. (2025). Bitcoin USD (BTC-USD) price history. Retrieved January 15, 2025, from https: //finance.yahoo.com/quote/BTC-USD/history

[8] Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on Loess. Journal of Official Statistics, 6(1), 3-73.

[9] Bouri, E., Gkillas, K., Gupta, R., & Pierdzioch, C. (2021). Forecasting realized volatility of Bitcoin. Physica A: Statistical Mechanics and its Applications, 575, 126021.

[10] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780.

[11] Puoti, F., Pittorino, F., & Roveri, M. (2025). Quantifying cryptocurrency unpredictability: A comprehensive study of complexity and forecasting. arXiv preprint arXiv: 2502.09079.