Price Oracle Accuracy Across Blockchains: A Measurement and Analysis

Robin Gansäuer*, Hichem Ben Aoun*, Jan Droll, and Hannes Hartenstein

KASTEL Security Research Labs
Karlsruhe Institute of Technology, Karlsruhe, Germany

Abstract. Decentralized finance depends on accurate price oracles to ensure reliable smart contract operations. Chainlink, a leading decentralized oracle network, bridges blockchains with off-chain data via price feeds. This study quantifies Chainlink's accuracy by comparing its price feeds of eight networks to centralized exchanges for two trading pairs, using 50 million data points for MAPE analysis and descriptive statistics. The results reveal performance variability: for example, Polygon achieves high accuracy, while Ethereum and ZKsync exhibit greater deviations under market volatility. The study discusses trade-offs between threshold and heartbeat configurations, emphasizing their complementary roles in balancing accuracy and efficiency. In addition, arbitrage implications and inefficiencies are identified.

Keywords: Price Oracle · Accuracy · Blockchain · DeFi · Chainlink.

1 Introduction

The rapid evolution of decentralized finance (DeFi) underscores the pivotal role of accurate off-chain data in enabling reliable smart contract operations. Price oracles, which bridge on-chain systems and external data sources, facilitate autonomous smart contract execution without dependence on centralized authorities. Such decentralized systems enhance the security and transparency underpinning DeFi platforms. However, inaccuracies in price data pose significant risks, including arbitrage opportunities, market manipulation, and the malfunctioning of automated financial services.

Despite their foundational importance, price oracles remain understudied in a comprehensive, cross-platform context. Prior research has largely been confined to isolated blockchain environments or specific exchanges, leaving gaps in understanding their performance across diverse ecosystems. Centralized exchanges (CEXs), known for their high liquidity and consistent pricing, are widely regarded as reliable benchmarks for financial data. By contrast, Decentralized Oracle Networks (DONs), such as Chainlink, aggregate off-chain data from multiple sources and deliver it on-chain via mechanisms like data feeds. Nevertheless, their capacity to ensure precise and consistent price information across heterogeneous blockchains remains insufficiently explored.

* These authors contributed equally to this work.

To bridge these gaps, this study systematically examines the accuracy of Chainlink's price feeds by comparing them against CEX benchmarks and evaluating the impact of parameters such as blockchain selection, deviation thresholds, and heartbeat intervals.

The investigation is guided by two key research questions:

- RQ1: *How do blockchain selection, deviation thresholds, and heartbeat intervals affect price updates of Chainlink Price Feeds?*
- RQ2: *How accurately do Chainlink Price Feeds reflect off-chain price data, based on benchmarks from selected CEXs across blockchains?*

Addressing these questions, the study compares Chainlink's price feeds with data from two CEXs, evaluating their precision and consistency across a variety of blockchain networks using data from the full month of December 2024. While Chainlink serves as the focal point, the findings contribute to a broader understanding of price oracle functionality and accuracy within the DeFi ecosystem. By addressing both the diversity of platforms and the operational mechanisms of oracles, this research provides insight into the critical role of decentralized oracles in ensuring the stability and security of DeFi platforms.

The structure of this paper is as follows: *Section 2* provides basic information on price oracles and decentralized oracle networks at the example of Chainlink. *Section 3* reviews related work and identifies key research gaps. *Section 4* outlines the methodological framework designed to address the research questions. *Section 5* presents the empirical findings, while *Section 6* discusses their broader implications, study limitations, and directions for future work.

2 Background on Price Oracles: Chainlink

Chainlink offers a notable approach to overcome the *oracle problem* [3], a fundamental challenge in blockchain technology where smart contracts are unable to interact directly with off-chain data in a trustworthy fashion. These contracts, which enforce agreements based on predefined conditions, operate within a blockchain's environment. However, blockchains are intentionally isolated from external data to preserve their security and decentralization [5]. This isolation creates a challenge for smart contracts that require real-world information, such as financial market prices, weather data, or supply chain information, to function correctly.

To address this issue, Chainlink proposes a *Decentralized Oracle Network* that connects blockchains to external data sources. By establishing this connection, smart contracts are empowered to execute transactions incorporating real-world data, while striving to maintain the intended decentralization of blockchain systems. The core of Chainlink's system relies on *Data Feeds*, which aggregate and deliver data from multiple independent oracles, intending to ensure data integrity and availability. These features are particularly crucial for DeFi applications, such as lending platforms, synthetic asset protocols and decentralized exchanges [5].

2.1 Architecture

Chainlink operates through a dual-layer architecture comprising *off-chain* and *on-chain* components that work in close tandem to enable the transmission of external data to smart contracts. This layered approach is designed to reduce the risks of erroneous or malicious data interference [3, 5].

The off-chain component is driven by a decentralized network of *Chainlink nodes*, which fetch data from external sources, perform computations and transmit the results to the blockchain. This decentralized structure seeks to reduce reliance on single points of failure and to enhance resilience against potential adversarial actions. Once processed, the aggregated data is relayed via *on-chain aggregator contracts* for smart contract execution [5].

The on-chain component enhances data reliability by evaluating oracle performance, matching smart contract data requests with suitable nodes, and consolidating inputs from multiple oracles into a single output [5].

2.2 Data Feeds, Thresholds, Heartbeats, and Rounds

A key service of Chainlink is its *Data Feeds*, which aggregate and provide real-world price data for various assets. These feeds are relevant for DeFi applications that rely on real-world pricing information to execute smart contracts. They operate on a *push-based design*, where updates are pushed to the blockchain based on specific conditions. Updates are triggered by the following conditions:[1]

- Deviation thresholds, in short thresholds: Update triggered by specified price change compared to the price currently reported on-chain. These thresholds are intended to ensure prompt reflection of market movements, enabling smart contracts to respond.
- Heartbeats: Periodic updates ensure data reliability during low market volatility, preventing outdated pricing from impacting smart contracts.

When a threshold is reached and a corresponding Chainlink node detects it, a new round is initiated to aggregate oracle data of the Chainlink nodes. Rounds are numbered, and therefore, one can associate reported prices to rounds. Values for thresholds and heartbeats need to be configured to provide timely and relevant data while optimizing network load and transaction costs. In Subsection 5.1, we analyze deviation thresholds and heartbeat intervals for various blockchains, while in Subsection 5.2, we analyze price accuracy, and in Subsection 5.3, we relate threshold and heartbeat configuration to price accuracy [2].

[1] See https://docs.chain.link/architecture-overview/architecture-decentralized-model

3 Related Work

3.1 State of the Art

Research on Chainlink Price Feeds is limited but critical due to their foundational role in DeFi. Key studies by Nadler et al. (2023) [11], Vakhmyanin and Volkovich (2023) [12], and Gogol et al. (2024) [7] provide valuable insights.

Nadler et al. (2023) conducted an extensive study of Chainlink oracles on Ethereum, analyzing over 150 million data points from 40 price feeds. Their investigation highlighted how heartbeat intervals and deviation thresholds impact oracle price accuracy. While effective in stable conditions, these mechanisms caused significant delays during periods of market volatility, exposing smart contracts to stale and unreliable data. Their update types and rounds structure analysis is foundational but focuses solely on the Ethereum blockchain [11].

Vakhmyanin and Volkovich (2023) explored Layer-2 ecosystems, focusing on the Mycelium platform on Arbitrum. They found that Chainlink's 2.5% deviation threshold delayed updates during volatile conditions, creating arbitrage opportunities for high-frequency traders. The study emphasized the difficulty of maintaining oracle accuracy in fast-changing environments [12].

Gogol et al. (2024) analyzed arbitrage opportunities across Layer-2 networks, including Arbitrum, Optimism, Base, and ZKsync. They identified over 500,000 unexploited arbitrage events and revealed how gas fees and block production times shape price discrepancies. However, their work does not address Chainlink's specific configurations or its potential role in reducing inefficiencies [7].

3.2 Research Gaps

While significant progress has been made in understanding Chainlink oracles in DeFi, key gaps remain.

First, Nadler et al. (2023) focus on Ethereum-based oracles, leaving the performance of Chainlink Price Feeds on other ecosystems, such as Layer-1 blockchains like Avalanche and Layer-2 solutions like Arbitrum, largely unexplored. These ecosystems introduce varying conditions, such as differing block times and transaction costs, that could significantly impact oracle behavior and require a broader cross-blockchain evaluation [11].

Second, while studies have examined mechanisms like heartbeat intervals and deviation thresholds, their effectiveness across different blockchain environments and under varying market conditions remains insufficiently explored. Gogol et al. (2024) address arbitrage dynamics, highlighting the influence of cross-rollup conditions and market variability, but do not analyze how these mechanisms directly affect Chainlink's accuracy or reliability in practice [7].

Third, no rigorous comparison has been made between Chainlink and CEXs such as Coinbase and Kraken across multiple blockchains. While prior research notes deviations during market volatility, the extent of Chainlink's accuracy and latency compared to CEXs remains unquantified, particularly under diverse and challenging conditions. Such comparisons are crucial for a thorough assessment of Chainlink's reliability and performance as a decentralized data provider.

4 Methods

This study adopts an adapted version of the *Knowledge Discovery in Databases (KDD)* process [6] to evaluate the accuracy of Chainlink price feeds compared to CEXs. Data was continuously collected for the entire month of December 2024 using a custom-built agent, capturing real-time price movements. This period was chosen due to practical constraints. The KDD framework comprises six stages: *Data Sourcing*, *Selection*, *Preprocessing*, *Transformation*, *Data Mining*, and *Interpretation and Evaluation*. To support this, the *Medallion Architecture* [4] organizes data into three layers: *bronze* (raw data), *silver* (cleaned data), and *gold* (refined dataset).

4.1 Data Sourcing and Selection

BTC/USD and ETH/USD trading pairs were chosen for their high liquidity, relevance in cryptocurrency markets, and availability across Chainlink and CEXs. Chainlink price data was accessed via public *Remote Procedure Call (RPC)* endpoints, capturing updates every second to ensure high-frequency data collection. For CEXs, websocket connections facilitated continuous, low-latency price streams. All data points were timestamped upon retrieval from the RPC, forming the *bronze layer*. The dataset included over 50 million data points.

Coinbase and Kraken were selected for their high market share and provision of free websocket access to USD-based trading pairs. We did not choose Binance because it does not offer trading pairs that include fiat currency, specifically USD. For Chainlink, price feeds spanned Layer 1 blockchains such as Ethereum, Avalanche, Binance Smart Chain, and Celo, alongside Layer 2 solutions including Arbitrum, Polygon, ZKsync, and Linea. These networks were chosen for their accessibility via public RPC endpoints, ensuring consistent and reliable data retrieval. To enhance reliability, we used a list of at least five RPCs as fallback options, with Llama RPC as the primary for consistency and Infura RPC as the final fallback.

4.2 Preprocessing and Transformation

The preprocessing phase ensured that raw data was systematically prepared for analysis. Missing values, occurring due to occasional delays in data updates, were handled conservatively to maintain data integrity. Temporal alignment across data sources was achieved by synchronizing timestamps to a unified one-second precision. These steps were executed using *Databricks* with *Apache Spark* on *Azure*, enabling large-scale data processing.

In the transformation phase, semi-structured data from the *bronze layer* was cleaned and converted into structured formats in the *silver layer* and further refined into an analysis-ready dataset in the *gold layer*. Data from Chainlink feeds and CEXs was aggregated to one-second intervals to enable consistent comparisons, ensuring alignment of update frequencies.

4.3 Data Mining

In the data mining phase, statistical methods were applied to assess the accuracy of Chainlink price feeds compared to CEX prices. Price accuracy was evaluated using the *Mean Absolute Percentage Error (MAPE)*, which quantified percentage-based deviations between Chainlink's price feeds and CEX prices from Coinbase, Kraken, and their weighted average. In addition, Pearson correlation coefficients are computed to quantify the impact of threshold and heartbeat parameters on price deviations.

4.4 Interpretation and Evaluation

The comparison framework, illustrated in Figure 1, is organized into two primary components: the *Analysis Component* and the *Comparison Component*.

The *Analysis Component* focuses on update types and round analysis, categorizing Chainlink price updates into mechanisms such as *threshold*, *heartbeat*, and *unknown*. This structured approach ensures a detailed and comprehensive understanding of the mechanisms driving Chainlink's price feeds.

The *Comparison Component* evaluates the accuracy of Chainlink price feeds for BTC/USD and ETH/USD against benchmark data from Coinbase and Kraken. Using *MAPE* and Pearson correlation, this component assesses deviations and examines how well Chainlink updates align with reference data under varying network configurations.

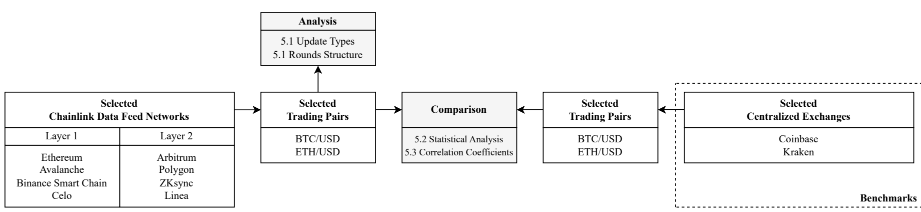

Fig. 1. Comparison Framework: Analyzing Chainlink Feeds Against CEX Benchmarks

Table 1. Chainlink Data Feed Network and Trading Pair Configuration

Layer 1 Networks				Layer 2 Networks			
Network	**Trading Pair**	**Threshold (%)**	**Heartbeat (s)**	**Network**	**Trading Pair**	**Threshold (%)**	**Heartbeat (s)**
Ethereum	BTC/USD	±0.5	3600	Arbitrum	BTC/USD	±0.05	86400
	ETH/USD	±0.5	3600		ETH/USD	±0.05	86400
Avax	BTC/USD	±0.1	86400	Polygon	BTC/USD	N/A	60
	ETH/USD	±0.1	86400		ETH/USD	N/A	60
BSC	BTC/USD	±0.1	60	ZKsync	BTC/USD	±0.5	86400
	ETH/USD	±0.1	60		ETH/USD	±0.5	86400
Celo	BTC/USD	±0.1	86400	Linea	BTC/USD	±0.5	86400
	ETH/USD	±0.1	86400		ETH/USD	±0.5	86400

5 Results

5.1 Analysis of Update Types and Rounds Structure

In this subsection, we examine how threshold and heartbeat configurations affect the temporal distribution of updates. Figure 1 shows the different data feed networks we focus on. Table 1 shows all analyzed data feeds with their configuration. The networks can be classified based on their update strategies: some use a hybrid approach, where both threshold and heartbeat are applied, while others follow a pure strategy, relying on only one of the two. Most networks, however, adopt the hybrid strategy, with significantly varying heartbeat intervals. As a result, we categorize the networks into three groups based on their heartbeat intervals: long, moderate, and short. The only network using a pure strategy is Polygon, which forms a fourth group.

For visual analysis, we make use of scatter plots as used in [11]: The plots show the price difference (deviation) of the update relative to the current round price on the vertical axis and the time since the last update normalized by the heartbeat interval on the horizontal axis. Indicated are the types of the updates: threshold, heartbeat, or unkown (for the case of an update that we cannot relate to threshold or heartbeat).

The first group is comprised of Arbitrum, ZKsync, Linea, Avax, and Celo. These networks make use of a long heartbeat interval of 86,400 seconds (24 hours). The deviation thresholds are set to $\pm 0.05\%$, $\pm 0.1\%$ or $\pm 0.5\%$. As shown in Figure 2 for the BTC/USD trading pair on ZKsync, updates occur predominantly close to the beginning of the heartbeat interval, triggered by significant market movements that exceed the threshold. Thus, a "long heartbeat interval" actually indicates reliance primarily on deviation threshold-driven updates.

The second group consists solely of Ethereum that adopts a hybrid mechanism with a moderate heartbeat interval of 3,600 seconds (one hour). In Ethereum, a wider deviation threshold of $\pm 0.5\%$ is used. The on-chain price data feed shows a more dispersed update pattern, as seen in Figure 3 for the BTC/USD trading pair. Updates occur inside and outside the threshold corridor, primarily influenced by threshold-driven updates. The temporal distribution of updates is wider and shows a more balanced distribution across the heartbeat interval. This balance reflects an approach of optimizing for a consistent update frequency while accommodating significant price movements to ensure robust performance in its decentralized price feed system.

The third group consists of BSC that incorporates short heartbeat intervals of 60 seconds and a price threshold of $\pm 0.1\%$. Figure 4 illustrates the hybrid update mechanism, where most updates are time-driven. This approach ensures frequent updates while responding to market fluctuations, providing adaptability to both stable and volatile conditions. The update distribution for BSC shows periodic peaks that do not correspond to its heartbeat interval, suggesting a periodicity within the heartbeat interval. This hybrid approach makes BSC particularly flexible in accommodating a range of market dynamics.

The fourth group is represented by Polygon, which relies exclusively on time-driven updates because of the absence of a defined deviation threshold. With a 60 second heartbeat interval, updates occur at regular intervals, as shown in Figure 5 for the BTC/USD trading pair. This ensures consistent data availability, albeit at the cost of efficiency during periods of market stability. Simliar to group three, the update distribution shows a periodicity that does not correspond to the heartbeat interval. Unlike threshold-driven networks, Polygon's mechanism sacrifices responsiveness to price deviations for the sake of predictable and consistent updates.

The diversity in update distributions across networks demonstrates the adaptability of Chainlink's system, balancing market responsiveness with consistent data availability. This flexibility allows the architecture to adapt to varying network requirements and market conditions. However, given the different update strategies, one is wondering whether all of these strategies lead to accurate price information. We address this aspect in the next subsections.

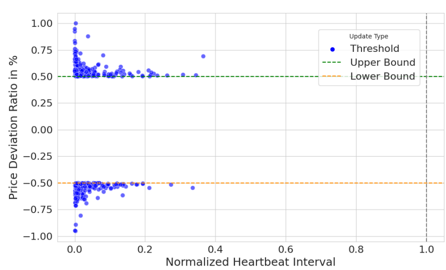

Fig. 2. Long heartbeat interval: updates of BTC/USD on ZKsync (threshold: ±0.5, heartbeat: 86,400s) for Dec. 2024

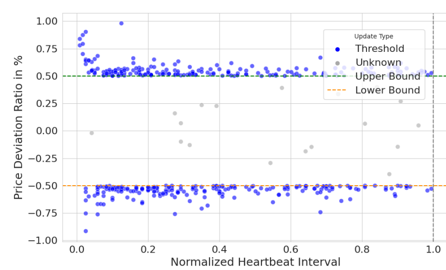

Fig. 3. Moderate heartbeat interval: updates of BTC/USD on Ethereum (threshold: ±0.5, heartbeat: 3,600s) for Dec. 2024

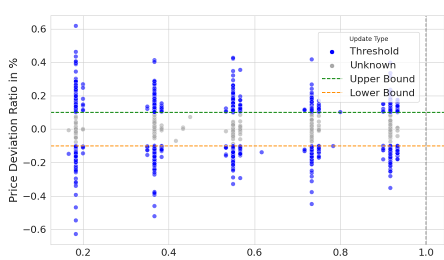

Fig. 4. Short heartbeat interval: updates of BTC/USD on BSC (threshold: ±0.1, heartbeat: 60s) for December 2024

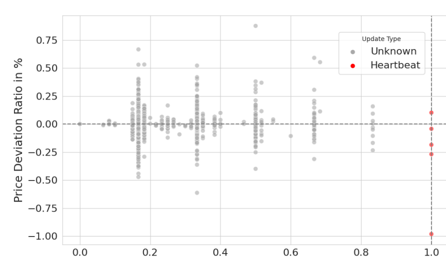

Fig. 5. Short heartbeat interval: updates of BTC/USD on Polygon (no threshold, heartbeat: 60s) for December 2024

5.2 Statistical Analysis of Price Accuracy

The statistical analysis in this section focuses on the accuracy of Chainlink's price feeds in replicating CEX price movements. Price movements should generally remain within the predefined upper and lower bounds that result from the configuration of the deviation threshold and the heartbeat interval. The reported price on Chainlink, the upper and lower bounds of the respective configuration, and the prices reported by Kraken and Coinbase are depicted in Figures 6, 7, 8 and 9 for ZKsync, Ethereum, Avax, and Polygon, respectively. The figures show data from a randomly chosen 12-hour window from 00:00 UTC until 12:00 UTC on 2024-12-03, as shorter intervals lack sufficient rounds and larger timeframes reduce the visibility of price deviations. The results illustrate varying levels of fidelity across networks, with notable discrepancies and fluctuations in some cases, while others demonstrate high accuracy and stability.

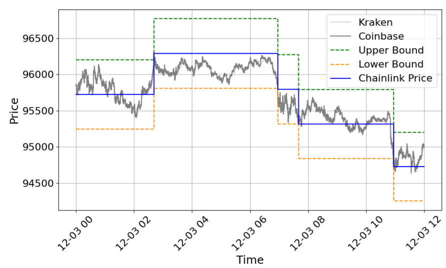

Fig. 6. Price and thresholds of BTC/USD on ZKsync (threshold: ±0.5, heartbeat: 86,400s) over time (12h)

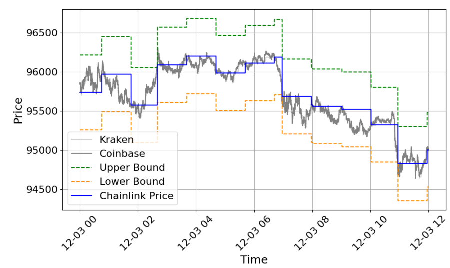

Fig. 7. Price and thresholds of BTC/USD on Ethereum (threshold: ±0.5, heartbeat: 3,600s) over time (12h)

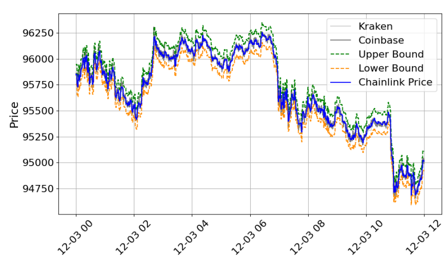

Fig. 8. Price and thresholds of BTC/USD on BSC (threshold: ±0.1, heartbeat: 60s) over time (12h)

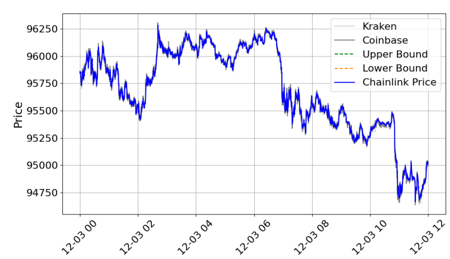

Fig. 9. Price and thresholds of BTC/USD on Polygon (no threshold, heartbeat: 60s) over time (12h)

Figure 6 reveals considerable price fluctuations on the ZKsync network. Centralized exchange prices frequently diverge between the upper and lower bounds, indicating significant volatility. While Chainlink's price feed effectively remains within the defined threshold, the wide range between the bounds illustrates the challenge of capturing precise market movements during volatile conditions. This observation is critical, as it highlights the limitations of relying solely on predefined bounds in such scenarios, where centralized exchanges exhibit inconsistent price behavior. The results suggest that while ZKsync's configuration ensures adherence to the threshold, it may struggle to accurately replicate real-time market prices during rapid fluctuations.

Similarly, Figure 7 depicts the behavior of Chainlink's price feeds on the Ethereum network. Configured with a wider threshold of $\pm 0.5\%$, the upper and lower bounds show substantial variability compared to the actual prices on Coinbase and Kraken. This variability, combined with frequent deviations between the bounds, suggests that the Ethereum configuration prioritizes update efficiency over strict adherence to real-time CEX prices. The Chainlink feed, while staying within the thresholds, captures a generalized trend rather than replicating precise market fluctuations. This trade-off between accuracy and stability is particularly evident in scenarios of increased market activity, where CEX prices reflect frequent and sharp changes.

In contrast, Figure 8 provides an intermediate perspective, showing Chainlink's performance on the BSC network. Although the price feed adheres to the threshold bounds, the deviation between the upper and lower bounds remains narrower than the deviation observed on ZKsync and Ethereum. This suggests a more balanced approach, where the Chainlink feed effectively captures real-time market trends while maintaining a reasonable level of stability. The results indicate that BSC's configuration balances accuracy and resilience.

Lastly, Figure 9 highlights the exceptional performance regarding accuracy of Chainlink's price feeds on the Polygon network. Despite the absence of a defined price threshold, the Chainlink prices align almost perfectly with those of Coinbase and Kraken, with minimal deviation. The graph demonstrates a seamless replication of the centralized exchange prices, suggesting that Chainlink effectively captures real-time market movements without significant lag or error. This precise alignment underscores the robustness of the Chainlink mechanism on Polygon, particularly in maintaining data accuracy even during periods of high market activity.

The histogram in Figure 10 presents a comparative analysis, using MAPE, based on the data of the full month of December, across multiple networks for three categories: Weighted CEXs (by transaction volume), Coinbase, and Kraken. The price deviation ratio was calculated by taking the absolute difference between the Chainlink price and the CEX price, divided by the Chainlink price. This resulted in a deviation value for each recorded data point. These values were subsequently averaged, grouped by network and trading pair to receive the MAPE. It reveals significant disparities in the accuracy of Chainlink's price feeds, with ZKsync exhibiting the highest MAPE across all categories, indicating substantial deviations from CEX prices. In contrast, BSC and Polygon consistently achieve the lowest MAPE values, demonstrating superior accuracy and alignment with market prices. Networks like Avax and Linea fall into a mid-

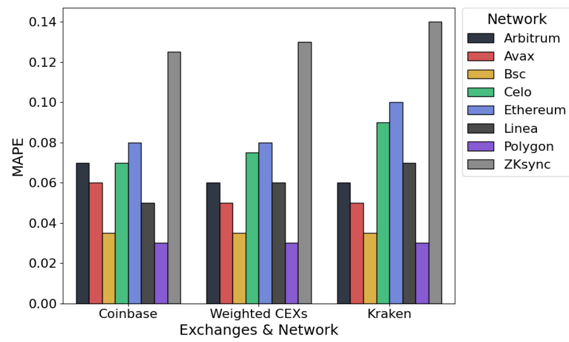

Fig. 10. MAPE for Accuracy Comparison by Exchange and Network

dle ground, showing moderate deviations. This analysis highlights the impact of network-specific configurations on price feed fidelity, emphasizing the need for tailored adjustments to improve performance on networks with higher errors.

Overall, the analysis reveals significant differences in the accuracy of Chainlink's price feeds across networks. Polygon demonstrates exceptional alignment with CEX prices, while AVAX strikes a balance between stability and accuracy. In contrast, Ethereum and ZKsync exhibit larger deviations, particularly during volatile conditions, reflecting the trade-offs of their configurations. The comparative MAPE analysis further underscores these disparities, highlighting Polygon and Arbitrum as the most accurate networks and ZKsync as the least.

5.3 Correlation of Update Parameters and Price Accuracy

To understand the influence of the configuration of the deviation threshold and the heartbeat interval on the price deviation ratio, we computed the corresponding Pearson correlation coefficients as descriptive statistics. The Pearson correlation coefficient was used to assess how changes in the deviation threshold and heartbeat interval influence the price deviation ratio, quantifying the strength and direction of the relationship between these parameters and the accuracy of the price data. The Pearson correlation coefficient for the threshold metric evaluates to 0.66 and indicates a strong positive relationship between the deviation threshold configuration and the discrepancy between the weighted price of the CEXs (weighted by transaction volume) and Chainlink's price feeds. This value suggests that as the threshold increases, the magnitude of price discrepancies also increases proportionally. This highlights how the threshold allows for greater divergence between reported and market prices before an update is triggered, leading to larger deviations as thresholds rise. For the heartbeat metric, the Pearson correlation coefficient is 0.27, which means a somewhat weak positive relationship with the weighted CEX price discrepancy. Please note again that we use the correlation coefficient as a descriptive statistic and refrain from inference.

6 Discussion

In December 2024, the observed mean absolute percentage error for the eight analyzed Chainlink data feed networks is smaller than 0.14 for all networks, and smaller than 0.1 for all observed networks but ZKsync. Thus, generally speaking, the average tracking error appears to be reasonably small. However, the update mechanism should react quickly on the one hand and should not induce unnecessary transactions on the other hand. In the following subsections we discuss this trade-off, further implications, limitations, and future work.

6.1 Tradeoffs

The analysis of MAPE values provides additional context on the accuracy of Chainlink price feeds across blockchains. Polygon stands out as the most accurate, achieving consistently low MAPE values. Notably, it operates with an average block time of just 2 seconds, significantly shorter than most other networks. This rapid block time facilitates frequent updates, enhancing alignment with off-chain prices despite a 60-second heartbeat and the absence of a deviation threshold. Such exceptional performance underscores the effectiveness of frequent, dynamic updates, raising the question of whether the threshold is always necessary. The answer depends on network-specific factors, such as block time and transaction costs. For Polygon, the absence of a threshold eliminates latency in corrective actions, but this approach may not be feasible on networks with higher costs or slower block times.

In contrast, ZKsync exhibits the highest MAPE values, reflecting significant deviations from off-chain prices. This is due to its higher threshold and longer heartbeat intervals. Avax and Linea occupy a middle ground, with moderate MAPE values indicating a balance between accuracy and efficiency. ARB also demonstrates high accuracy, performing comparably with Polygon, particularly in aligning with CEX prices.

The interplay between the threshold and heartbeat parameters highlights their distinct yet complementary roles, borrowing concepts of safety and liveness from the field of distributed computing. While the threshold acts as a safety condition, ensuring that updates are triggered only when deviations exceed a predefined limit ("nothing bad has happened") [8], the heartbeat serves as a liveness condition, guaranteeing that updates eventually occur to restore alignment ("eventually something good will happen") [1]. Removing the heartbeat would be problematic, as it ensures eventual recovery from discrepancies, even in low-volatility markets. Without it, price updates could stall for extended periods, causing outdated on-chain prices that increase the risk of incorrect liquidations and arbitrage exploits. Conversely, a poorly calibrated threshold can lead to prolonged inaccuracies, especially during high-volatility conditions. These principles are supported by the update patterns observed across networks. Blockchains with thresholds showed few updates triggered by the heartbeat, reinforcing its role as a safety net rather than a primary driver of updates. Without the heartbeat, there is no assurance that discrepancies will be corrected in a timely manner, highlighting its indispensability in maintaining data accuracy.

The tradeoff between oracle accuracy and operator costs is critical. Higher accuracy demands frequent updates, lower deviation thresholds, and shorter heartbeat intervals, increasing on-chain fees and computational expenses. Conversely, raising thresholds or extending heartbeats reduces costs but risks delayed updates and larger discrepancies between on-chain and off-chain values. While minimizing the accuracy threshold can significantly cut costs, eliminating it entirely would trigger updates for every minor price change, making operations economically unfeasible, especially for high transaction cost blockchains. These costs would ultimately shift to users.

To balance accuracy and cost, a dynamic threshold adjustment could provide a more effective solution. By adapting the threshold based on factors such as the rate of price change at the update trigger, market volatility, network congestion, and trading volume, oracles could optimize updates to reflect meaningful price shifts while avoiding unnecessary transactions.

6.2 Implications for Practice and Research

From a practice perspective, the findings highlight a potential vulnerability in DeFi systems, where price deviations, even as low as an average of 0.15%, can lead to substantial arbitrage profits in scenarios involving high trading volumes. For instance, if Chainlink's BTC/USD price feed deviates on average by 0.15% compared to CEX prices, an arbitrageur could exploit this by purchasing BTC on a DeFi platform at the lower oracle price and immediately selling it at the higher market price after an update. With significant trading volumes (e.g., $10 million), a deviation of 0.15% translates into a profit of $15,000 per arbitrage cycle, compounded across multiple transactions. Empirical evidence confirms that such inefficiencies are actively exploited in practice. Recent research indicates that about 52% of Ethereum blocks contain at least one flashbots transaction, with 98.68% of liquidations on AAVE and Compound explicitly depending on Chainlink oracle updates occurring within the same block [9]. In addition, approximately 85% of these bundles consist precisely of one oracle update immediately followed by a liquidation, illustrating traders' already existing strategic exploitation of oracle update timing to maximize profitability [9]. Such practices potentially result in liquidity providers incurring losses as their provided liquidity is depleted to fund trades executed at suboptimal prices. This systematic exploitation underscores the necessity for refining oracle mechanisms, such as optimizing deviation thresholds and heartbeat intervals. Yet, even when such inefficiencies are detected, their exploitability can be limited by transaction costs, slippage, and protocol configurations like liquidation penalties, highlighting the need to assess the practical feasibility of these solutions at the ecosystem level [10].

From a research perspective, these findings underscore the need to examine how oracle discrepancies impact the broader DeFi ecosystem, particularly in terms of financial burden distribution between stakeholders such as liquidity providers, governance participants and end-users. Quantifying these losses under varying conditions provides critical information on the stability of the protocol. Adaptive oracle mechanisms, capable of dynamically adjusting thresholds and update intervals based on metrics like market volatility and trading volume, also warrant evaluation for their role in mitigating price deviations effectively.

Furthermore, the systemic implications of oracle discrepancies, including their effects on liquidity pools, derivatives pricing, and AMM strategies, highlight potential vulnerabilities that propagate losses across interconnected protocols. The economic and ethical dimensions of arbitrage require scrutiny, focusing on whether arbitrage profits enhance capital efficiency or destabilize protocols. The potential shift from push-based data feeds to pull-based data streams also raises questions about trade-offs in latency, coordination, and computational costs, offering a critical perspective on optimizing decentralized price feed systems.

6.3 Limitations and Future Work

Our research presents several limitations and opportunities for future work. The data collection for one month, conducted during the moderately volatile period of December 2024, provides valuable information but may not fully capture the impacts of extreme market conditions, such as high volatility spikes, severe congestion, or periods of low activity. Extending the study over varying market cycles would provide a more holistic understanding of oracle performance and resilience. The analysis was limited to two trading pairs, both involving a fiat currency (BTC/USD and ETH/USD), which, while significant, may not reflect the unique dynamics of crypto-to-crypto trading pairs. Including a broader set of trading pairs, such as ETH-BTC or stablecoin-to-crypto pairs, could reveal additional complexities and patterns in price feed behavior. In addition, reliance on CEXs for price comparisons does not fully account for deviations from DEX prices, which may result from differing liquidity profiles, slippage, or arbitrage activities. Further analysis could quantify these deviations and their implications for oracle reliability under diverse market conditions.

Additionally, the use of indexed data to analyze network trends, while effective for identifying overall patterns, may obscure more granular network-specific behaviors, such as the impact of block time variability on price update latency and accuracy. Networks with shorter or highly variable block times could exhibit distinct update characteristics that warrant closer examination. Price deviations between DEXs and CEXs, driven by factors such as liquidity fragmentation, slippage, and differing trading mechanisms, also require detailed investigation. Quantifying these discrepancies would provide valuable insights into how varying market structures influence oracle reliability.

7 Conclusion

This study provides an evaluation of the accuracy of Chainlink's price oracles across diverse blockchain networks. By comparing price feeds with data from centralized exchanges, the findings reveal how threshold and heartbeat parameters interact to balance price accuracy with update frequency. The analysis examines different approaches of oracle configurations. For instance, Polygon's time-driven updates achieve high accuracy with minimal deviations, while larger threshold values on ZKsync and Ethereum present challenges under volatility. Avax exemplifies a balanced approach, maintaining stability while reflecting real-time market trends. These findings underscore Chainlink's architectural adaptability across diverse market and network conditions.

Our findings identify potential vulnerabilities in DeFi systems that rely on push-based data feeds. Even small deviations in oracle accuracy might lead to arbitrage opportunities and systemic inefficiencies. Whether these vulnerabilities are actually exploitable is subject to further research. However, adaptive oracle mechanisms capable of dynamically tuning parameters to match market volatility and the adoption of pull-based data streams could significantly improve resilience.

References

1. Alpern, B., Schneider, F.B.: Recognizing safety and liveness. Distributed Computing **2**(3), 117–126 (1987)
2. Breidenbach, L., Cachin, C., Chan, B., Coventry, A., Ellis, S., Juels, A., Koushanfar, F., Miller, A., Magauran, B., Moroz, D., et al.: Chainlink 2.0: Next steps in the evolution of decentralized oracle networks. Chainlink Labs **1**, 1–136 (2021)
3. Caldarelli, G.: Understanding the blockchain oracle problem: A call for action. Information **11**(11), 509 (2020). https://doi.org/10.3390/info11110509
4. Databricks: What is the Medallion architecture? (2020), https://docs.databricks.com/en/lakehouse/medallion.html, accessed: 2025-01-17
5. Ellis, S., Juels, A., Nazarov, S.: Chainlink: A decentralized oracle network. Whitepaper (2017)
6. Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: From data mining to knowledge discovery in databases. AI Magazine **17**(3), 37–54 (1996). https://doi.org/10.1609/aimag.v17i3.1230
7. Gogol, K., Messias, J., Miori, D., Tessone, C., Livshits, B.: Cross-rollup MEV: Non-atomic arbitrage across l2 blockchains. arXiv preprint arXiv:2406.02172 (2024)
8. Lamport, L.: Proving the correctness of multiprocess programs. IEEE Transactions on Software Engineering **SE-3**(2), 125–143 (1977). https://doi.org/10.1109/TSE.1977.229904
9. Messias, J., Pahari, V., Chandrasekaran, B., Gummadi, K.P., Loiseau, P.: Dissecting Bitcoin and Ethereum transactions: On the lack of transaction contention and prioritization transparency in blockchains. In: Financial Cryptography and Data Security. pp. 221–240 (2024). https://doi.org/10.1007/978-3-031-47751-5_13
10. Muck, M., Schmidl, T., Wolf, J.: Wish or reality? On the exploitability of triangular arbitrage in cryptocurrency markets. Finance Research Letters **73**, 106508 (2025). https://doi.org/10.1016/j.frl.2024.106508
11. Nadler, M., Schuler, K., Schär, F.: Blockchain price oracles: Accuracy and violation recovery. Available at SSRN (2023)
12. Vakhmyanin, I., Volkovich, Y.: Price arbitrage for DeFi derivatives. In: Proceedings of the IEEE International Conference on Blockchain and Cryptocurrency (ICBC). pp. 1–4 (2023). https://doi.org/10.1109/ICBC56567.2023.10174884