Precog Interval Forecast Impact in Simulated Trading

Uriel Morone

May 2025

Abstract

Subnet 55, Precog, tasks Bittensor miners with modeling two outputs: a point forecast and an interval forecast of the Bitcoin Price. In a previous article, we showed that Precog Bitcoin price point forecasts contained valuable information that consistently outperformed holding Bitcoin in different market environments. We repeat this analysis for a new 40 day window from April 1 through May 10, 2025 and add an adjustment to vary the trade size depending on the width of the interval forecast. We show that not only does the original strategy maintain reliable improvement over a buy-and-hold strategy, but that the new variable size trading strategy often shows improvement over the original constant trade size approach. We test over 18 different sample periods of 14 days and find the constant trade strategy beats holding in 9 samples and loses in 5, while the variable-trade strategy beats holding in 12 samples and loses in only 2. Finally, we conclude with some statistical metrics measuring the forecasts which we will use to benchmark as our subnets evolve

Keywords: Bittensor, Bitcoin, Forecast, Prediction, Trading

1. Introduction

This article is designed to demonstrate the value inherent in the Interval Forecast estimates produced by miners on Precog, Subnet 55. This serves as a followup to a previous article, *Evaluating Precog Subnet 55 Forecast Performance via Simulated Trading*, published April 8th 2025. The methods are generally similar: simulating a basic low-complexity trading strategy that makes decisions entirely based on the Precog outputs and current Bitcoin price. Here we extend the method to incorporate information from the Interval Forecast, in addition to the Point Forecast used previously. Additionally, we include some statistical metrics to benchmark performance in the future and illuminate the relationship between predictions and observed changes in price.

Although it is relatively easy to understand, the interval forecast captures more sophisticated dynamics than the point forecast. It incorporates information about the variance of the price in a way that anyone without mathematical training can understand. Unlike simple variance, in principle the interval can be asymmetrical, biased in one direction if the market is in the midst of a run for example. This makes it a more flexible and nuanced measure than a simple standard deviation. An interval is evaluated with a maximum possible score when its upper and lower bounds exactly match the max and min price observed during the time period considered.

Unlike the point forecast we covered previously, the interval prediction consists of two numbers and covers an extended time period. The point forecast is measured against a single target value, the price one hour in the future. However the interval forecast is evaluated against all price values available during the ensuing hour; at the 1-second frequency used this means 3600 data points are used in the evaluation. This more sophisticated evaluation reflects the additional dynamics considered by the miners when coming up with a forecast.

The exact methodology is covered in detail in our documentation, along with illustrated examples. In general terms, the interval forecast is evaluated by the product of a *width*-factor and *inclusion*-factor. The former penalizes intervals the wider they get. It has a max value of 1.0 when the predicted interval spans the range of the observed prices, and decreases to a value of 0.0 as the interval becomes infinitely wide. The inclusion-factor is designed to punish predictions which fail to actually capture all of the observed prices. The inclusion-factor hits a max value of 1.0 when all observed prices lie between the lower and upper interval bounds, and has a value of 0.0 when the interval fails to include any observed price at all.

In the studies that follow, we simulate trades that implicitly use the interval forecasts as a measure of *confidence*. This is not necessarily the only way, or indeed even the best way, to apply the interval forecasts but we show that nevertheless there is some valuable information contained therein. As always our analysis never intended prescriptive but rather open-ended, showing one possibility with the intention of inspiring further applications by the eventual end-user.

1.1 Caveats

In the interest of transparency, we mention a few caveats and potential shortcomings to the following analysis. There's many potential knobs to turn, hyperparameters so to speak, in the trading strategies considered and even in how we arrive at the predictions themselves. At any time there are roughly 240 miners submitting predictions, and we know for a fact these are not of equal quality. Determining how many miners to consider, and how to incorporate them can impact results. This goes even more so because the exact configuration of miners change, and even those that remain may adjust their models over time. We cannot guarantee that the subnet will serve as a precise machine with churns exactly reproducible results.

In our previous article we used the average predictions of the top 20 miners. However, since that article was published, we experienced a surge in miner churn mirroring the increase in our α-token price. This in turn meant that a larger proportion of miners were using default strategies from the base miner class we provide. For this reason we reduced the number of "top" miners used to determine the "subnet forecast" from 20 in the previous article to 5 in the current article. This showed improved results which were sufficient to demonstrate the value of the predictions. We recognize that there is some subjectivity here and that the optimal number of predictions may fluctuate over time depending on subnet dynamics.

Since observing the increased interest following the previous article, we made a few changes to improve the integrity of the forecasts and reward the most valuable miners. We reduced the AdjustmentAlpha subnet hyperparameter, accelerating the rate at which registration fee increases with additional demand. This ensures new miners are paying market rate, and have less incentive to hoard slots opportunistically. More importantly, we changed the subnet mechanism to decrease incentives dramatically in the case of many tied miners. This scenario is especially relevant for many miners running the base miner code. In addition to significantly increasing the share of rewards going to participants proposing intelligent predictions, it also lowers the value of "playing it safe" with one's predictions and never proposing something that could end up too-wrongly.

Because a subnet is a living thing in some sense, we are always looking for ways to incrementally improve collective performance. The end result is that our metrics will always be attempting to evaluate a moving target.

2. Investigation

As mentioned above, there's many possible applications of the interval prediction. It could be used to trade volatility index futures, calculate an asymmetrical hedging strategy, or as an estimate of confidence in current market prices. Our hope is that by forecasting a value with an obvious interpretation, eventual end-users will be able to easily determine how to transform it and incorporate it into broad range of models with different targets. This analysis considers the latter possibility, assuming the interval forecast can serve as a measure of confidence in future bitcoin price movements.

The methodology uses a benchmark "range", determined by looking the hourly maximum - minimum BTC price in 1 hour windows for the last several weeks. This value can fluctuate, but was in the range of $545 and this value was used as our benchmark parameter. When the subnet interval forecast was larger than this value we scaled down the size of the trade, with an eye towards "high risk". When the interval was narrower than this benchmark we increased the size of the trade proportionally. The trade size was clipped at 0.2x and 5x the base line trade.

2.1 Methodology

To analyze the value of the interval forecasts we used the following simulated trading strategy, then compared it to holding bitcoin over different time frames:

First, to condense each individual miner to a "Forecast" for both the point and interval predictions, we considered average rewards given to each miner by 4 approved validators (RoundTable21, Rizzo, Yuma, and OTF). At each prediction time we filtered the top *five* miners by rewards. The value of each prediction type (point, lower interval, upper interval) is averaged over the top 5 miners at each time step to give a single point forecast and interval forecast at each time.

Then, the trading strategy follows this simple algorithm:

1. Assume there are both **USD** and **BTC** wallet balances. In this excercise, these can *only* be positive.
2. At each prediction time if the point prediction is greater than the current Bitcoin price then "buy": reduce the USD balance by `[trade volume]` amount of USD and increase the BTC balance by `[trade volume]*[BTC Price]` BTC.
   - The size of the trade is determined by the relative size of the interval to the a previously observed baseline: $vol_{trade} = 545.0/(upper_{intv} - lower_{intv})$. Where the trade volume (vol_{trade}) was given bounds 0.2 and 5.0 to avoid erratic edge cases, and $upper_{intv}$ and $lower_{intv}$ are the bounds of the interval forecast.
   - If the point prediction is less than the latest observed price, do the opposite: increase USD balance by `trade volume` and reduce BTC balance (ie. *sell*) an equivalent amount.
3. Repeat this process every 5 minutes, at each prediction time.

2.2 40 Day Period

We first consider the entire time period since our previous article, to understand how the strategy evolves over a longer period and see in detail the comparison of our new variable-size trading strategy to the previous fixed-size method.

A trading simulation like this has many parameters and assumptions that can be modified, but in an attempt to make it as clear as possible we settled on the following rules:

- Every strategy is limited by its total capital. If a trader does not have access to USD they cannot buy more BTC, and if they do not have BTC they cannot short it for USD. This is in contrast to some parts of our previous article where we assumed unbounded liquidity.
- Every strategy starts with the same capital, in the same allocation. In this case that's $100 of value, allocated entirely to BTC.

Fig 1 shows the evolution of three different portfolios, all of which are initialized with the same starting balance.

1. Fixed Size: A binary trading strategy that buys or sells $1.00 based solely on the point forecasts
2. Variable Size: A fluctuating trading strategy that determines the buy-or-sell decision via point forecasts, but determines the magnitude of the trade using the interval forecast (as described in Section 2.1)
3. Hold BTC: A baseline strategy that assumes the portfolio commits all capital initially to BTC and never sells the initial amount.

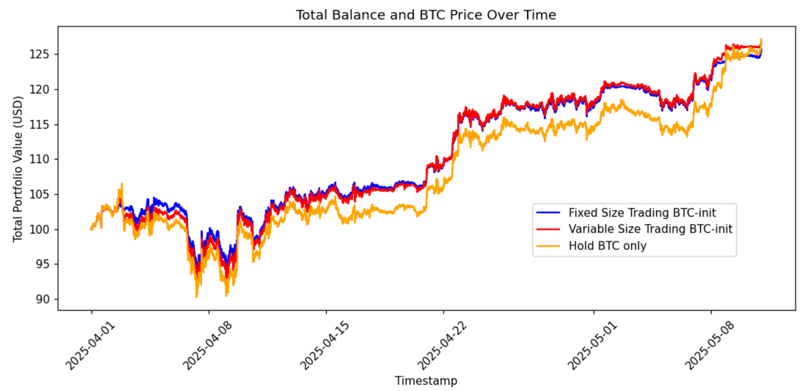

Figure 1: 40-day Performance

What we observe is that both Precog-based trading strategies meaningfully outperform pure Bitcoin performance for the *majority* of the time period considered. It is important to note that if we look at the total results the difference is minor and, in fact, not actually positive for the trading strategies. While this may be disappointing to someone actually trading, it is important to be clear on the challenge we have posed ourselves. From 4/01 to 5/10 the Bitcoin price increased almost 27%. It is extraordinarily difficult for any strategy to outperform fully allocating to an asset that grows that quickly. Precisely speaking as of 05-11 00:00 the original \$100 of BTC were worth \$126.99, while the fixed-size and variable-size strategies had \$125.68 and \$126.59 USD of value respectively.

However, it is important to note that the 40-day window we consider here is arbitrary. We could have just as easily selected 35 days or 45 days, if we had a bit more time. Each of these choices would change the final performance. Rather than look at the final accumulated value, it is informative to consider what percentage of the time the strategies outperform the holding. From the figure we immediately see over almost the entire window both trading strategies outperform, until essentially the last day. Inspecting the results we see the that the fixed-size strategy outperforms holding during **92%** of the window, while the variable-size strategy outperforms over **96%**. Another way to think about this is if any random time you decided to cash out before May 11, its overwhelmingly likely you would have profited more with a trading strategy.

Although in this bigger picture the variable-strategy approach did not result in a massive difference, the story changes a bit when we bootstrap several shorter time samples later in Section 2.3. Performance is always going to depend on the specific time period considered and the more samples we can look at the better.

Nevertheless, Figure 2 can help us understand how the variable-sized strategy is performing and what it says about our interval predictions. We see here two plots showing the relative balance in the USD and BTC wallet for each of the two trading strategies. Often, the variable-size strategy plays things safer. One reason it did not outperform the fixed-size strategy even more is because the variable strategy did not sell as aggressively in the first few days, when BTC dropped. At other times, such as the period from 4/18 to 4/20 the variable strategy was more aggressive.

This is a good spot to note that the precise *mechanism* we used to incorporate the variable trading strategy may be far from optimal. Its possible there's better ways of setting the baseline parameter which corresponds to a trade of $1.00. Its also possible there is a better way to vary the size than proportional scaling. As before, our goal is not to find the best trading strategy but to find *something* that shows value, even if imperfect.

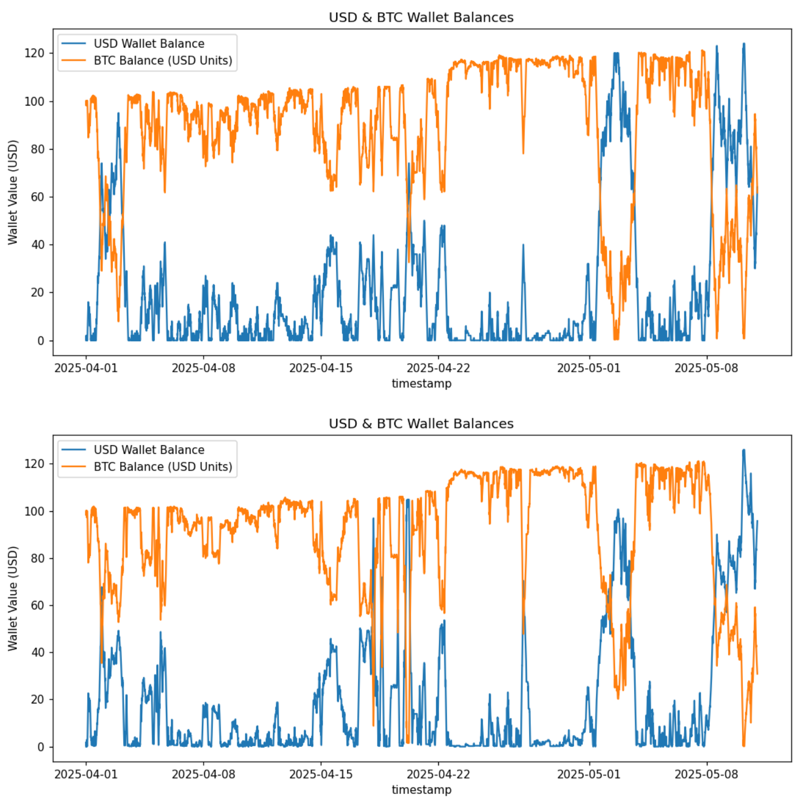

Figure 2: 40-day Simulation Wallet Balances. Top: Fixed Trading Strategy. Bottom: Variable Trading Strategy

2.3 Bootstrapping 14-day samples

To control for the influence of specific dates, we broke our 40 day window into 18 samples. The start times range from 4/01 00:00 to 4/26 12:00. Each sample began 36 hours apart, for a total of 18 starts, and looked at performance over the following 14 days. The full table of data is reported in Appendix A, but we summarize the results here.

If we start by considering the final net worth of each portfolio we find that both strategies outperform holding in a majority of scenarios, and the variable strategy shows a significant improvement in consistency. For these purposes the final value is counted as a tie when the portfolio values are within $0.50. The fixed-size strategy "loses" in 5 of the 18 scenarios, while the variable-size strategy only underperforms in 2. The variable strategy results in 6 times as many wins as losses.

Outcome	Fixed	Variable
Ties	4	4
Wins	9	12
Losses	5	2

Table 1: Comparison of Trading Strategies Final Portfolio Value to Buy-and-Hold. Ties occur when the strategy's total USD value was within $0.50 of Holding approach.

Like the previous section, we can also consider the *percentage of time* each strategy outperformed during the interval, rather than the final value. Coincidentally, the "losses" look similar, 4 and 2 for the fixed- and variable- strategies respectively. However a close look at Apx A will reveal that the samples with a net loss relative to holding were different than the samples where the strategy underperformed holding for a majority of the time. This just shows how challenging it can be to quantify these types of performances

Outcome	Fixed	Variable
Ties	3	2
Wins	11	14
Losses	4	2

Table 2: Comparison of Outcomes when looking at Percent of Time with greater portfolio value than Holding strategy. Ties were considered the range 40-60%.

When looking at the percent of time when each strategy did better than holding, the results are actually rather overwhelming. A majority of the "wins" come in the form of samples with over 90% of the window outperforming the hold strategy. In other words, these strategies are not particularly volatile. When they outperform Holding, they do so almost the whole time. Only the two samples on 04-19 and 04-20 showed scenarios when the strategy mostly under-performed, but even these samples recovered in terms of net worth right at the end.

These samples show a more obvious value add from the interval predictions. The losses are halved, and two losses and one tie are converted to clear "wins" when going from the fixed-size strategy to a variable-size strategy. These three samples represent 16.6% of the 18 samples considered.

3. Statistical metrics

Having demonstrated value in a simulation, we conclude this article with some more quantitative statistical metrics. Below we calculate correlations of the point forecast with the hourly returns of the BTC price and the average success rates of the interval forecasts. Without context these numbers do not mean much, but we intend to start benchmarking and hope to see improvements as we refine miner incentives and continue improving the subnet design.

3.1 Correlation of Point Forecast with Hourly Return

This section analyzes the Point Forecast calculates its correlation with hourly returns. The correlations are not enormous, although they are positive. Because of the number of predictions (12 every hour, 288 a day) stacked over several days we see the performance above. We hope this value will continue to improve as miners get better, and Precog sees success attracting more sophisticated forecast models.

The correlations were calculated by looking at the relative prediction at each time stamp and comparing to the hourly return one hourly later. Relative prediction in this case means the point prediction value minus the Reference Rate at time of prediction. The hourly return meanwhile is the percent change in the Reference Rate from the prediction time to the evaluation time. These values are shown visually in a scatter plot in Figure 3.

We calculate two correlations coefficients: Pearson and Spearman. The Pearson Correlation the most commonly used and measures the linear relationship between two variables, where -1.0 and 1.0 represent perfect lines with negative and positive slopes. The Spearman Correlation measures monotonicity, or the likelihood that both variables increase together and decrease together, which makes it more suitable for non-linear relationships.

Our forecasts have Pearson and Spearman correlations of 0.009 and 0.03. These are fairly small and indeed we can see in Fig 3 the scatter plot is quite noisy. However after many time points we're able to extract profitable information. We hope to improve these numbers as the best miners rise to the top and we accumulate longer times series to understand subnet outputs.

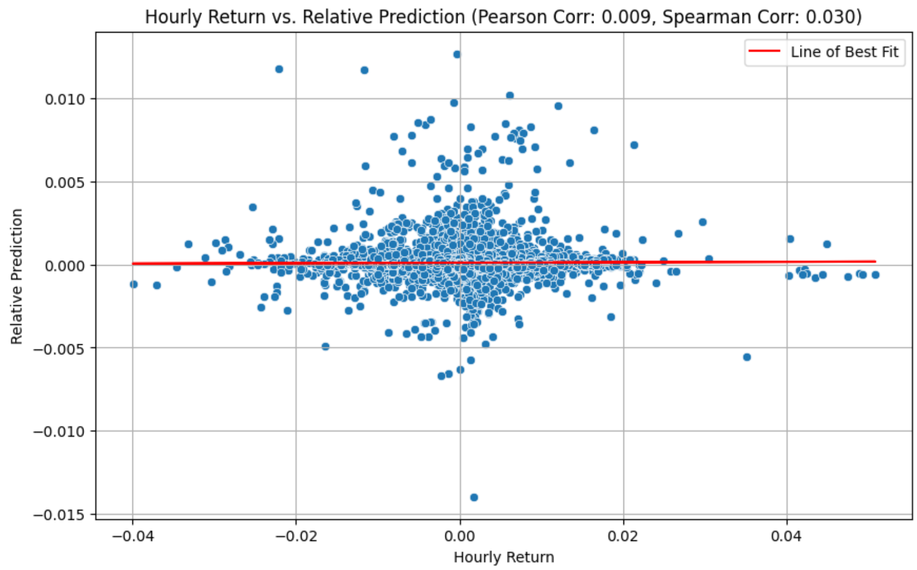

Figure 3: Scatter Plot of Relative Prediction vs Hourly Return

3.2 Interval Forecast Sizing

A simple benchmark to help understand the interval forecast is what percentage of prices currently lie within the Interval. In our methodology this is evaluated by the *inclusion-factor* term. At each time we consider the real prices for the next hour and simply calculate how many lie between the interval predicted at the initial time.

What we see is a little bit surprising. Overwhelmingly the intervals include the vast majority of prices. In some sense this is good, as the interval forecast can set a reliable upper bound on the variability of prices. However, we do want the intervals to be as tight as they can be while capturing the range. The stats show (Table 3) that more than half of the predictions include 100% of the price points, suggesting there may not be sufficient incentive to reduce the interval size.

This is reinforced when we look at the relative width. Relative width here means the range of the interval forecast divided by the min-max range of observed prices during the period. We see in Table 3 the typical interval is more than 2 times the actual observed price range. The bright side is this suggest there is room for improvement in making the interval forecast more precise. We intend to review the incentive and make sure we are directing miners to produce the most valuable outputs possible.

Statistic	Percent Included	Relative Width
Mean	87%	2.52
Min	0%	0.20
25% Quartile	83%	1.45
50% (Median)	100%	2.16
75% Quartile	100%	3.24
Max	100%	22.72

Table 3: Percent of Price Points Included and Relative Width

4. Conclusion

We extended our previous analysis to show the intelligence of the Precog Interval Forecasts by using them in a simulated trading strategy, which modifies the trade volume according to the forecasts. Our findings show that application of the interval forecast can improve the reliability of a Precog trading strategy significantly. The "win-loss" ratio, compared against a baseline of holding the asset increased from 11-4 to 14-2, when moving from the previous fixed-size to the new variable-size trading methodology. Moreover, this past month reinforces our previous findings that the Point Forecast itself contains profitable information which, while not exclusively positive, can consistently improve returns over holding Bitcoin.

Finally, we share some mathematical evaluation benchmarks, which is an area we hope to continue developing in order to inform the community and future end-users about our subnet performance. Our analysis of interval forecasts suggest they may be wider than necessary, and we plan to dig more deeply to understand why and whether the incentive mechanism to reward the miners needs to be tuned.

A. 14-day Sample Data

Start	Final Value Fixed	Final Value Var	Final Value Hold	% Times Fix > Hold	% Times Var > Hold	% Times Var > Fix	Net Change BTC Price
04-01 00:00	104.95	104.57	102.49	89%	89%	7%	2056.02
04-02 12:00	101.83	101.81	99.38	98%	98%	19%	-526.19
04-04 00:00	103.43	103.96	102.18	99%	96%	60%	1810.84
04-05 12:00	103.03	103.83	102.14	51%	94%	85%	1785.88
04-07 00:00	109.77	110.73	108.69	57%	99%	82%	6811.21
04-08 12:00	110.29	111.61	110.84	41%	58%	92%	8670.44
04-10 00:00	113.80	114.47	113.50	73%	83%	88%	11146.18
04-11 12:00	115.66	116.42	115.28	77%	90%	98%	12545.07
04-13 00:00	111.83	112.21	111.01	100%	99%	54%	9392.36
04-14 12:00	113.10	113.77	112.30	99%	97%	68%	10445.99
04-16 00:00	112.38	113.12	112.72	29%	56%	86%	10641.29
04-17 12:00	113.06	113.83	113.77	15%	61%	98%	11643.54
04-19 00:00	113.63	114.31	114.75	5%	34%	87%	12458.53
04-20 12:00	113.27	113.45	113.44	1%	3%	56%	11304.52
04-22 00:00	108.78	108.95	108.25	65%	66%	70%	7221.18
04-23 12:00	104.53	104.76	103.83	83%	90%	85%	3582.78
04-25 00:00	106.77	107.43	109.82	89%	92%	89%	9231.53
04-26 12:00	106.39	107.07	109.82	72%	79%	97%	9252.98