

HMX 1.75 June 17. 2026 Market Intelligence & Accuracy Report

Released by Heatmup Oy In June 8. 2026. The document covers the intelligence and accuracy of the underlying HMX 1.75 model. The performance and accuracy figures are explained but we have to acknowledge that the figures are currently validated by us, the creators of this model. We are happy to send the raw forecast history data, to any entity that wants to publicly validate our results. Heatmup is currently testing many models and thus this model may not be available in the near future. The forecasts are meant for informative purposes and do not constitute financial advice.

I. Executive Summary

This is the first accuracy report for HMX 1.75. It covers only the model's live, public period: roughly the last three months, every resolved forecast in that window, graded against what the market actually did. It is not a historical backtest. The full history study follows the release of HMX 2.0. HMX 1.75 is the unweighted baseline. Every prediction carries equal weight. Nothing is filtered, nothing is down-weighted, no model is trusted more than another. That is deliberate. It sets the floor. Whatever signal shows up here shows up with one hand tied, before any of the weighting that 2.0 introduces.

Each resolved forecast point is placed into the percentile band the actual price fell into, then tallied across every asset and date in the window. The actual price is the OHLC4 mid of the resolving bar. The model orders its percentiles correctly, builds a sensibly shaped distribution, and keeps its tails where tails belong. That is the evidence that HMX 1.75 is reading signal, not emitting noise.

They also agree that, in this window, realized outcomes sat above the model's central forecast. That much is in the data. Whether it is the model leaning or the market's direction over three months is the open question, addressed next. Both are published here with equal weight, because that is the rule.

Two explanations fit, and three months cannot separate them.

The first is the window. Calibration is a long-run property. Over any short stretch where the market mostly moved one way, even a perfectly calibrated model reads as shifted, because the sample is too small and too one-directional to average out. Three months is short. If the period leaned in one direction, that alone can produce the curve above, and it corrects itself as more regimes resolve.

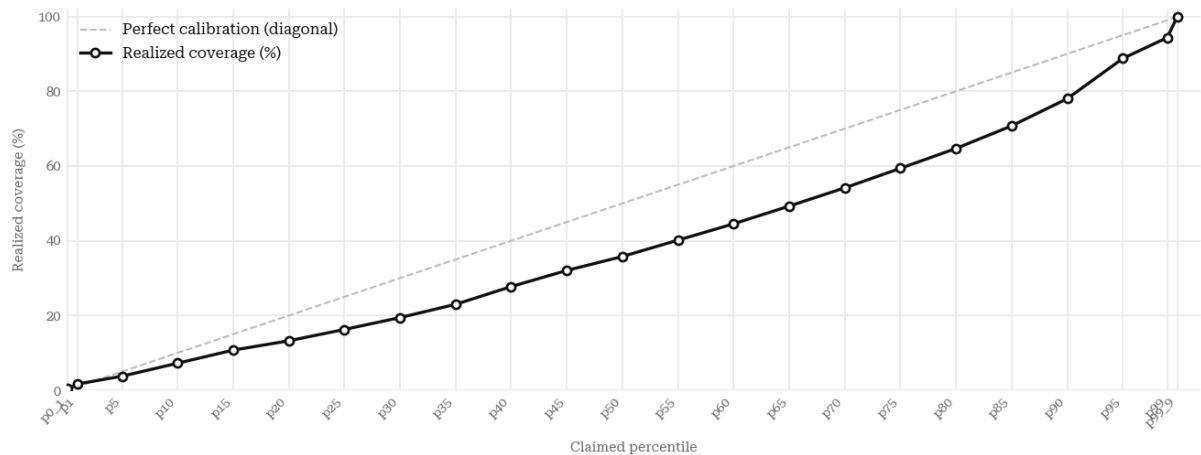
The second is the weighting. When every prediction votes once and nothing is filtered, a subset of models leaning one direction is never damped, and their tilt survives aggregation intact. An unweighted ensemble cannot correct its own location, and a standing bias of this kind would look the same on the chart.

These are not mutually exclusive, and the data here cannot tell them apart yet. That is the honest reading. HMX 2.0, later this year, weights predictions by their resolved track record and filters the ones that earn no trust, which removes the second cause if it is present. The first resolves itself with time and a longer record. This report is the baseline both will be measured against.

II. Model Calibration

Model Calibration: Claimed vs Realized

Fraction of outcomes at or below each percentile line



HMX 2.0 • Model Calibration • 03/06/2026

Claimed percentile against realized coverage. The dashed diagonal is perfect calibration.

Calibration is where the baseline shows its signal and its flaw at once.

Read the diagonal as truth: if the model says a price has a 50 percent chance of landing below a line, then over many forecasts half of them should. The model's curve tracks that diagonal in shape but sits below it through the whole middle of the range. At the median, only about 40 percent of outcomes landed below the line the model called the median.

Translated: realized prices ran above the model's center more often than they should have. The deviation is systematic, not random scatter, every percentile leans the same way. The curve is also smooth and monotonic, every percentile correctly ordered, which is the first thing that separates a real distribution from noise. So the shift is real in the sample. Whether it is the model or the window is a separate question, taken up below.

The drift varies in severity across the domain. The gap between the perfect calibration diagonal and the realized coverage reaches its maximum width in the upper-middle percentiles. At the p75 mark, realized coverage sits near 60 percent. The engine systematically underprices the probability of outcomes in this specific quadrant.

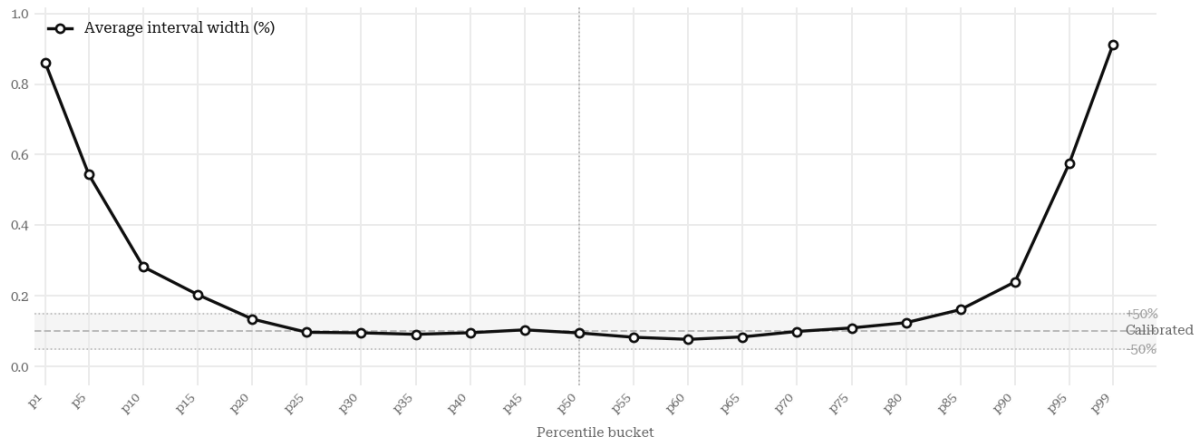
The calibration error compresses at the outer boundaries. At p5, realized coverage rests near 4 percent. At p95, it reaches 89 percent, and p99 hits 95 percent. The model bounds the extreme limits of the market accurately. It captures the absolute floor and ceiling but misplaces the distribution of mass inside that total range.

The unweighted architecture of HMX 1.75 explains the mechanics of this shift. The engine aggregates multiple independent lenses, such as options flow, macro risk, and news sentiment, and assigns them equal influence. A systematic miss in the central percentiles indicates a specific subset of these lenses maintained a bearish bias during the three-month live period. Because every prediction votes once and nothing is filtered out, a cluster of models projecting a lower median drags the aggregate center down without breaking the overall geometric shape of the curve.

III. Shape: the distribution is sound

Model Accuracy: Interval Width

Mean bandwidth variance across percentile thresholds



HMX 1.75 • Model Accuracy Report • 03/06/2026

Spacing between neighboring percentile lines, as a fraction, across the range.

This forms a clean smile: tight through the center, where most probability mass should sit, and wide in the tails, where it should. A noise generator does not produce this. It is the signature of a proper single-peaked distribution, and the baseline produces one consistently. The center of the smile sits inside the calibrated band, and the tails widen the way real tails do.

Between the p25 and p75 thresholds, the average interval width stays flat near 0.1. This plateau rests securely inside the calibrated tolerance zone. The expansion outside this core is sharp and uniform. The width reaches 0.85 at the p1 threshold and scales to 0.9 at p99.

This geometry is a direct product of the ensemble architecture. The engine combines independent lenses like options flow, prediction-market pricing, and macro risk models. These distinct standpoints naturally disagree on the likelihood of extreme tail events, which generates the necessary width at the edges. They agree far more closely on where the core probability mass belongs, which forces the center of the distribution to compress correctly. The system models the structure of the uncertainty rather than drawing a single momentum curve.

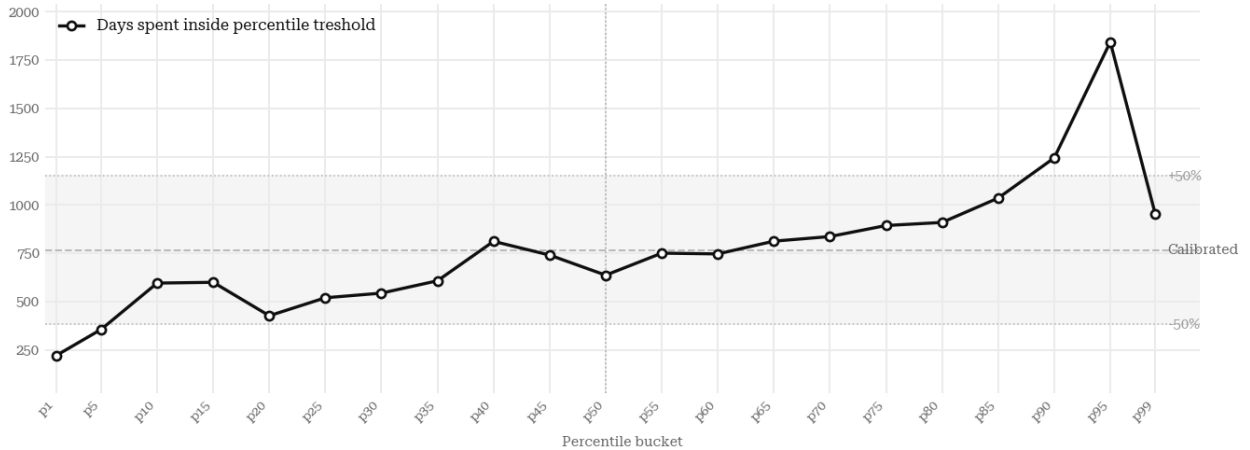
HMX 2.0 will introduce performance weighting to target much sharper, more accurate predictions. The upcoming engine filters out untrusted lenses and scales the influence of individual models based on their resolved track records. This active weighting mechanism, along with the foundational model calibration established in the 1.75 baseline, operates as the leading factor for those accuracy gains. The system retains the structural integrity of the current distribution while correcting the central drift, aligning the core probability mass with where the market actually resolves.

A perfect forecasting engine does not produce a probability distribution. It produces a single point. On the interval width chart, a perfect engine would drop the entire curve to zero. Every percentile boundary would collapse onto the exact same price. The model would state that there is zero variance between the 1st percentile and the 99th percentile outcome. That is the definition of absolute certainty driving up the growth of Heatmup every day.

III. Consistency: where the outcomes landed

Model Consistency: Time in Percentile

Retention duration across percentile thresholds



HMX 1.75 • Model Accuracy Report • 03/06/2026

Count of resolved outcomes in each percentile bucket. The flat reference is even, calibrated spread.

A perfectly calibrated model spreads outcomes evenly across the buckets. The baseline does not. Occupancy rises toward the upper percentiles and piles up around p90 to p95. This is the calibration chart seen from another angle. Outcomes clustered above the model's center. The buckets that should have been quiet, the high ones, did most of the work.

The calibrated baseline for this evaluation cycle sits near 750 days. The lower percentiles see significantly less retention. The p1 bucket holds slightly over 200 days, well below the -50 percent calibration bound. Retention climbs unevenly through the lower half of the distribution, dipping near 400 days at p20 before crossing the calibrated midline near p40.

The upper half of the curve demonstrates the systematic upward shift. From p60 onward, the days spent inside each threshold climb continuously. The p90 bucket registers roughly 1250 days, crossing the +50 percent error bound. The peak occurs at p95, which captured over 1800 days of collective outcome time. The extreme right tail at p99 drops back near 1000 days. The engine mapped the absolute outer edge of the price action, but the market spent vastly more time near the p95 ceiling than the unweighted baseline predicted.

V: Read it for what it is

These numbers are a first signal, not a verdict. The formal history study, across a longer record and more market regimes, comes with HMX 2.0.

The above results have been published, validated and calculated by Heatmup Oy.

Heatmup Oy CEO, Veikko Ahonen, 15/06/2026