

the WiseBot

Live • paper trading

WISEBOT • VERSION 3

Systematic

An honest, boring edge.

Third in a series documenting every version of WiseBot. This is the current version — the one running now. Unlike the first two, the honest verdict here isn't "it failed." It's something more modest, and more durable, than the win I originally went looking for.

the WiseBot

the-wisebot.com
June 2026

This is a living document. v3 is what WiseBot runs today, so this paper will be updated as the record grows. The numbers below come from backtesting and the start of live operation; the up-to-date, verifiable track record lives on the site — see the closing section. **This revision drops the last of the secrecy: the full strategy, the exact parameters, and the carry mechanics are now published in the open (see *The exact recipe*).**

The two failures, one rule left standing

Two versions in, here's the scoreboard. WiseBot v1 won 90% of its bets and lost money to a 1:24 payoff. WiseBot v2 moved to SOL perpetuals with a healthy 1:1 payoff and still lost — three strategies, all negative over a 14-day paper run, beaten by a single dull market regime. (Both stories are their own papers.)

The two failures left exactly one rule standing, and it's the foundation of everything here:

Prove the idea against years of history, across many regimes, before trading a cent — and prefer a robust result over a perfect one.

v2's deepest mistake wasn't a bad strategy; it was trying to validate one by living through two weeks of market in real time. v3 starts by refusing that. Before any live trading, the idea gets tested against ~5.6 years of data covering the 2021 bull, the 2022 bear, the 2023 recovery, the 2024—25 bull, and the 2026 drawdown — hundreds of regimes seen in minutes, not one crawling past in real time.

The thesis change

I threw out scalping entirely. Fast trading was a losing game for my setup: fees ate the edge, I had no speed advantage against professionals, and frequent trading in chop just bled money one tiny loss at a time.

v3 is the opposite kind of system: **systematic trend-following on a small basket of major crypto** (BTC, ETH, SOL), holding positions for days or weeks, not minutes.

In plain terms, trend-following means: when an asset is in a sustained uptrend, hold it long; when it's in a sustained downtrend, go short (bet on it falling); when it's just chopping sideways with no trend, stand aside in cash. It's one of the most-studied, most boring strategies in finance — the workhorse of managed-futures funds for decades. It is not exotic, and that's the point. The ability to go *short* and to go *flat* is also what gives it a shot at the one thing I actually want: not losing — even making a little — when the market falls.

Design choices you can follow

A few deliberate choices, each meant to keep the system simple and survivable:

- **~1x leverage.** Exposure roughly equal to capital. On a perpetual future this is the conservative setting: it all but removes liquidation risk (it'd take a ~100% move against me) and keeps financing costs minimal. I use perpetuals only because they're how you short — not to gamble with borrowed size.
- **Daily cadence.** Signals are checked a few times a day; decisions and rebalancing happen once a day. No 15-second loops. Low frequency means low fees, less noise, easier validation, and far less that can break.
- **Hyperliquid as the venue.** It carries the dominant share of perpetual-DEX volume, has low fees, a bot-friendly API, and a non-custodial "agent wallet" (a key that can trade but cannot withdraw). It also lists real perpetuals on gold, oil, and equity indices — which matters for where this is going (see *The ceiling*).
- **An LLM as risk officer — brake only, never accelerator.** Once a day, an AI layer reads macro and news and is allowed to do exactly one kind of thing: *reduce* risk — cut exposure, force flat, raise cash. It can never open a trade, add exposure, or push the system toward more risk. The reason is strict: an LLM is non-deterministic, can't be backtested, and can hallucinate, so it must never touch the part that makes money. But it can see things rules can't — an exchange hack, a regulatory ban, a macro shock — and that's worth capturing, but only as an emergency brake, with every intervention logged and reversible.

How I tested it honestly

The whole point of v3 is the validation method, so here's how it was actually checked — including the ways I tried to fool myself and didn't let myself.

- **Multi-year, multi-regime backtest** on ~5.6 years of BTC/ETH/SOL data, with a realistic cost model: real exchange fees, slippage, and actual historical funding (what you pay or earn to hold a perpetual). No fake, frictionless costs.
- **A 192-combination parameter sweep**, judged on the *distribution* of results, not the single best one. The question wasn't "what's the best backtest I can produce?" — that's how you overfit. It was "does a randomly chosen sensible setting still do fine?" The answer was yes: the median configuration beat a simple buy-and-hold on a risk-adjusted basis, and 37% of all combinations beat it outright. A wide good region, smooth gradients, no knife-edge — the opposite of an overfit.
- **Out-of-sample and walk-forward testing**, which is where most strategies quietly die. And here's the most important honest finding I have: **re-optimizing the parameters for each new period actively destroyed the edge.** Chasing the best recent settings is overfitting by another name. The system only held up when I fixed sensible parameters once and refused to touch them.

That last point is the discipline that runs through the whole project. The number-one risk isn't a bad market; it's me over-tuning until a backtest looks beautiful and means nothing.

What the numbers actually say

Here's the honest result, with no spin. Out-of-sample, with fixed sensible parameters and no re-optimization (Aug 2022 — Jun 2026):

	Risk-adjusted return (Sharpe)	Annual return	Volatility	Max drawdown
WiseBot v3 (trend sleeve)	0.93	20.2%	22.5%	-20.9%
Buy-and-hold (equal-weight)	0.62	21.2%	60.5%	-58.7%

Read that carefully, because the headline is not "it makes more money." It doesn't. Over this window it earned about the same as simply buying and holding the three coins (roughly 2x your money either way).

The edge is risk reduction, not return. Same return, but with about *one-third* the volatility and *one-third* the drawdown. Buy-and-hold got there by riding a -94% crash on the way; the strategy got there on a far smoother ride, with a maximum loss of about -21% instead of -59%. And critically — it earned money in the downturns that gutted buy-and-hold: roughly +1.6% through the 2022 bear and +4.1% through the 2026 drawdown, while holding lost 30–90%. That "boring curve that doesn't crater when the market does" is exactly the profile I was after.

I want to be blunt about what this is *not*: it is not a way to beat the market in absolute terms. If a long bull run is all you'll ever face, buying and holding makes more. The value here is for anyone who can't stomach — or survive — a 90% drawdown.

The ceiling, named

I'll name the limit plainly instead of hiding it. Three major crypto assets are highly correlated: in a sell-off they crash together, and the diversification you think you have collapses toward a single bet. I tested longer horizons and other variations to push past this; none added a *return* edge. The conclusion is that the ceiling is **structural, not a matter of a better signal**. Long/short among three things that move together gives you a risk edge, not a return edge.

The way out — and it's future work, honestly flagged — is adding genuinely uncorrelated legs: real perpetuals on gold, oil, and equity indices, which Hyperliquid supports on the same venue. That's the next major step, not something I'm claiming today.

The second sleeve: market-neutral carry

The trend strategy, on its own, is a risk-reduction engine — it doesn't reliably beat buy-and-hold on return. So the book pairs it with a **second, market-neutral sleeve** that harvests a *structural* premium these markets pay most of the time: the perpetual **funding rate**.

Here's the mechanism, in full. On a perpetual future, longs and shorts exchange a small payment roughly every hour — the *funding rate* — that tethers the perp price to spot. In crypto, most of the time, longs pay shorts: structurally more people want leveraged long exposure than short. You can collect that payment without taking a directional bet by going **delta-neutral** — buy the asset on the spot market and short the

same notional on the perpetual. The two price exposures cancel (if the coin doubles or halves, the long and the short offset), and what's left is the funding you earn for holding the short perp. I run it as an equal-weight basket across the same three coins: long spot, short perp, on each.

Its shape is the opposite of the trend sleeve's: it earns steadily when leverage demand is high, roughly nothing when markets are dead calm, and it rarely loses — *except* in the one situation named in the projections below, when funding flips negative and the sleeve pays instead of earns. Crucially, it is almost completely uncorrelated with the trend sleeve.

That lack of correlation is the whole point. WiseBot v3 is therefore a **two-sleeve book**, not a single strategy: a directional sleeve that protects in downturns, plus a market-neutral carry sleeve that earns in the quiet and the froth, blended at a fixed weight and rebalanced daily. They live in different regimes and cover for each other, which is why the blended book holds a better risk-adjusted profile than either piece alone.

The exact recipe

Earlier versions of this paper kept the parameters and the carry mechanics private — "transparency on results, opacity on method." I've dropped that line entirely. The wallet is public and on-chain, so anyone watching the trades can reconstruct the strategy anyway; guarding the recipe was security theater, and it sat badly in a project whose whole pitch is honesty. The edge here was never secrecy — it's discipline, low cost, and not over-tuning. So here is the entire machine, enough to rebuild it.

Trend sleeve, per asset, once a day:

- **Direction.** Long if price is above its 50-day moving average *and* its 120-day return is positive; short if below the average *and* the 120-day return is negative; otherwise flat. (The moving-average filter does the work of keeping it out of chop, so the momentum deadband is set to zero.)
- **Sizing.** Inverse-volatility: each asset is scaled by its own 30-day volatility so each contributes a similar share of risk, targeting **30% annualized portfolio volatility**, with **gross leverage capped at 1x**.
- **Rebalancing.** A 5-percentage-point no-trade band: a position isn't touched unless its target weight has drifted by more than that. This is what keeps turnover — and fees — low.

Carry sleeve: equal-weight long-spot / short-perp on BTC/ETH/SOL, exactly as described above.

The blend. Capital is split at a fixed weight of **0.6 to carry, 0.4 to trend**, rebalanced daily. Per asset, the net target works out to $\text{perp} = 0.4 \times (\text{trend signal}) - 0.6/3$ (always net short by the carry leg) and $\text{spot} = 0.6/3$ (the carry's long leg). That weight is deliberately *not* optimized — the projections below show why tuning it is a trap.

Costs assumed everywhere: the real Hyperliquid taker fee of 0.045% per trade — which I've since confirmed against live fills, it is exactly that — plus a conservative 2 basis points of slippage, and actual historical funding. Nothing frictionless.

What it would have done: starting one, two, three years ago

A single backtest window can hide how much the result depended on *when* you happened to start. So I ran the stricter test a real launch actually faces: pick a start date, choose the parameters using **only the data available before that date** (no peeking at the future), freeze them, and run the book forward to today — then repeat for each of the last several years. This is the honest version of "what if I'd turned it on back then."

Full two-sleeve book, production settings, costs as above, forward to the June 2026 data cutoff, against an equal-weight buy-and-hold of the same three coins over the identical window:

If launched...	Held for	WiseBot v3 (full book)	Max drawdown	\$1,000 →	Buy-and-hold basket	its drawdown
June 2025	1 year	+14.9%/yr (Sharpe 1.56)	-5.6%	\$1,151	-33.9%/yr	-59.7%
June 2024	2 years	+8.6%/yr (Sharpe 0.98)	-5.6%	\$1,181	-16.7%/yr	-57.9%
June 2023	3 years	+19.6%/yr (Sharpe 1.98)	-5.6%	\$1,709	+35.7%/yr	-62.1%

What each launch year would have returned

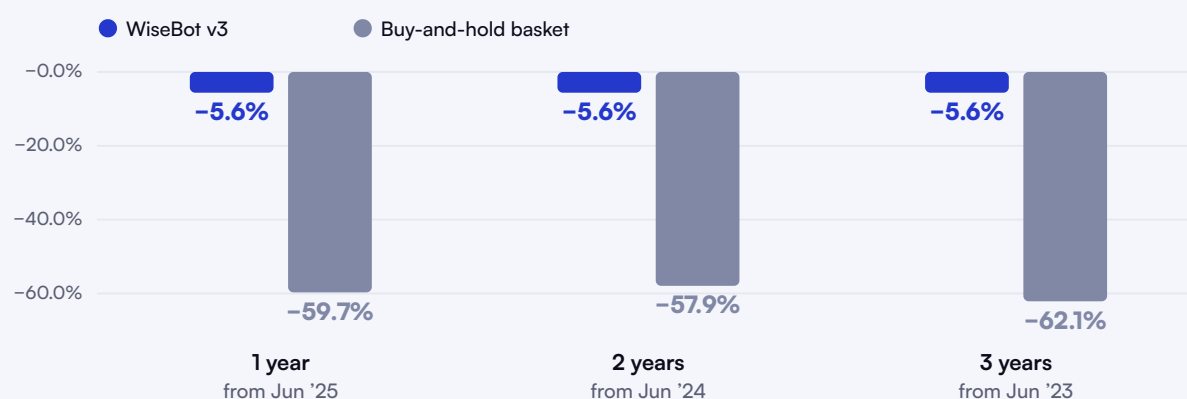
Annualized return — full two-sleeve book vs holding the basket, identical window.



The bot earned a positive return in every cohort; buying and holding lost money in two of the three. The exception is the bull window (from mid-2023), where holding made more in raw terms — the trade-off the bot makes on purpose.

The worst drop you'd have sat through

Maximum peak-to-trough loss in each window — closer to zero is better.



The whole point of the project in one picture: a worst-case fall of about -6% versus about -60% for simply holding — same market, same window, a different ride entirely. That -6% calm is regime-dependent, though; see the caveats above.

The pattern matches everything else in this paper. The book earned a positive double-digit return in every cohort, at a Sharpe of 1–2, with a worst drop of about **-6%** — while simply holding the coins meant riding a **-58% to -62%** drawdown every single time, and *losing* money outright in two of the three windows. In the one bull-heavy window (launching mid-2023), holding made more in raw terms (+35.7%/yr vs +19.6%/yr): the same trade-off as always — I give up upside in a roaring bull to never take the brutal drop.

For a longer horizon, the **trend sleeve alone** reaches back further — it only needs price history (available from 2020), whereas the carry sleeve's funding data begins in May 2023. Launched in June 2022, the trend sleeve returned +20.0%/yr at a -20.9% drawdown over four years (\$1,000 → \$2,072), versus +16.1%/yr at -59.9% for buy-and-hold.

Three things I won't hide about this table:

1. **That -5.6% drawdown is suspiciously tidy — and it's the carry sleeve's calm, which is regime-dependent.** None of these windows contained a true *funding stress*: a deleveraging cascade where funding flips negative and the carry sleeve pays instead of earns, exactly when you'd least want it. Funding does go negative in falling markets — it has, repeatedly, across this very history. I do not promise that -6% holds through a real funding shock; the honest expectation is a deeper drawdown the day one arrives.
2. **The carry weight (0.6) was never optimized.** Had I let the computer choose it to maximize the backtest, it would have gone all-in on carry and printed a Sharpe near 9 — a fantasy number manufactured by a calm window. I fixed the weight by judgment and refused to tune it. The discipline *is* the strategy.
3. **The production parameters benefit from being chosen on long history.** In the strictest test — re-deriving the parameters from only the short history available at each start date — the multi-year cohorts matched the production settings closely, but the one-year cohort, with too little history to choose from,

would have picked worse parameters and returned about +2% instead of +15%. I'm not claiming the settings were chosen blind; I'm showing both, so you can see the gap.

Every figure here is reproducible from the public backtest code (`npm run vintage`).

What I tried and threw away: timing the market's mood

The most natural objection to all of this is one I put to myself: *the bot gives up upside in a bull market — so why not detect the bull, simply hold through it, and only switch the short-capable machine back on when the trend turns down?* It sounds obviously right. That's exactly why it had to be tested before being believed — "obviously right" is the disguise most overfit rules wear.

So I built it and ran it the honest way, on the full history. The rule is simple: for any coin in a confirmed uptrend — price above its long-term moving average — hold it long, like a buy-and-holder; for everything else, run the normal trend logic that can short. One new dial, the length of that moving average, fixed in advance rather than tuned.

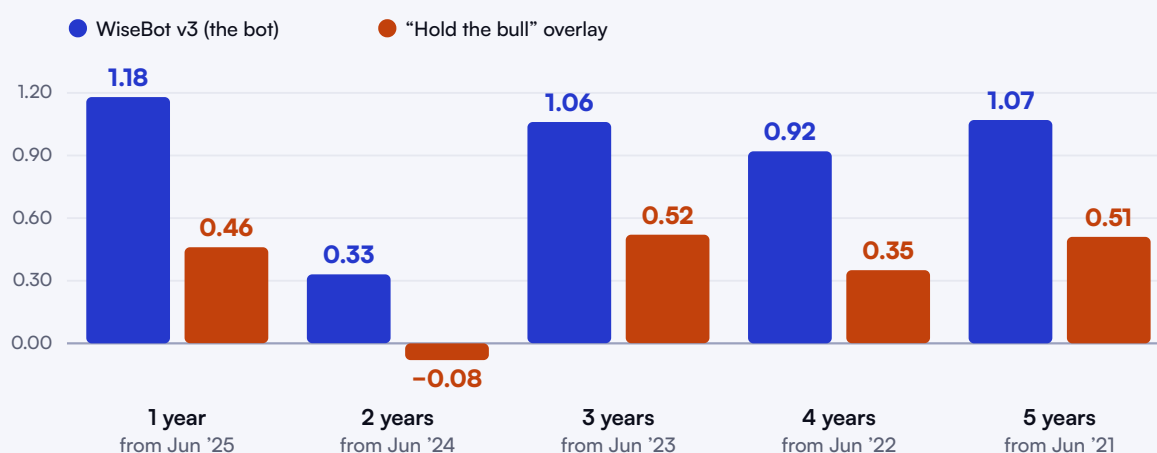
The results point in two opposite directions, and the contradiction *is* the lesson.

On the **full 5.5-year sample** — with the 2021 bull market in view — it looks like a triumph. With a 50-day filter the overlay returns about 82% a year at a Sharpe of 1.45: better than the bot, better than holding. A backtest to fall in love with. But look at two tells. First, the "fixed" dial isn't innocent: 50 days gives 82%/yr, 100 days gives 50%, 200 days gives 40%. A robust rule does not nearly double its return on a number I picked by hand. Second, almost all of that return is one bull market — the 2021 run — that this particular window happens to contain.

So I ran the test that removes hindsight: the same launch-cohort method as the section above. Pick a start date, freeze the rule, run it forward to today — repeated for each of the last five years, with the moving-average dial held at a single fixed value throughout.

The obvious idea, tested out of sample

Risk-adjusted return (Sharpe) by launch year — the bot vs a regime-timing overlay. Higher is better.



In every one of five launch years, "just hold through the bull" did worse than the plain bot — and at two to three times the drawdown. The cheapest place to kill a seductive idea is the backtest, not the account.

Out of sample, the verdict is unanimous and brutal: **in all five launch cohorts the regime overlay did worse than the plain bot** — a lower risk-adjusted return every single time, at a drawdown often two to three times deeper (around -55% versus -21% in the multi-year cohorts). It frequently failed to beat simple holding too. The "hold through the bull" rule reliably collected the *downside* of holding — the deep drawdown — without reliably catching the upside.

The reason is the one fact no timing rule escapes: a regime is only obvious in hindsight. By the time price has fallen back through its moving average to confirm the bull is over, you've already eaten half the crash. The overlay holds long straight into the top, lags the exit, and forfeits the protective short that earns the bot its keep when markets fall. The worst of both worlds — measured, not asserted. The whole experiment took an afternoon and cost nothing, which is precisely the point: the cheapest place to kill a seductive idea is in the backtest, not in the account. (Reproducible: `npm run regime .`)

This is also why the boring two-sleeve book stays as it is. The bot is *already* a regime detector — trend-following is nothing but leaning long in uptrends and short in downtrends — and pairing it with the market-neutral carry sleeve is the most honest way I've found to be smart across regimes without pretending to call the turns in advance.

Where the question stays genuinely open — and why there's an AI layer at all. A *hardcoded* timing rule fails; that much is now measured. Whether an *adaptive* one could do better is honestly unresolved, and it's part of why v3's design leans on a supervised AI risk layer rather than a fixed rulebook. Today that layer is strictly a brake, exactly as described earlier, and nothing in the validated edge depends on it. But it is also built to be an on-going, in-the-open experiment: a place to learn, in live operation, whether something that reads the market and the news each day can manage risk more adaptively than a static rule can — and, in time, whether that adaptive judgment can add to returns rather than only defend them. I'm flagging that last ambition as exactly what it is: unproven, and years from earned. It will only ever graduate from "brake" to

anything more if it first survives the same out-of-sample discipline that every number in this paper had to pass. Until then it brakes, every call is logged, and every call is reversible.

Where it stands now

As of writing, v3 has cleared its validation gate and is in **paper trading** — running the validated engine forward on live market data, on the real venue, with no real capital yet. If it holds up, a small amount of real money goes in, and the wallet becomes public. The LLM risk officer gets attached after that, in live operation, as the conservative brake described above — it is *not* part of the edge that was validated, and I won't pretend otherwise.

From the moment real capital is live, the commitment is fixed: a public, on-chain wallet anyone can inspect; the same honest metrics always shown — return, **max drawdown**, expectancy, risk-adjusted return, never just the flattering number; and signals time-stamped before the fact, so I can't claim a call after it's already paid off.

There's no asymmetry left here. I used to draw a line — total transparency on results, opacity on method — and I've erased it. The wallet is public and on-chain; the trades are visible; the method is inferable by anyone who cares to watch. Guarding the recipe was theater, and it sat badly next to a project whose entire pitch is honesty. So you get all of it: every number, a wallet you can verify, **and** the full strategy and parameters, laid out above. What makes the track record trustworthy was never the secrecy — it's that you can check it yourself.

This document will keep being updated as that record grows. **The current, verifiable track record — live equity curve, drawdown, the on-chain wallet — lives at the-wisebot.com**. If anything here ever disagrees with what the wallet shows on-chain, believe the wallet. That's the whole idea.

Methodology note

The analysis and drafting of this document are AI-assisted (I work with Claude as my stated method — the same AI also runs as the bot's risk officer, under the strict limits described above). Every backtest figure comes from the validation runs on real historical data with realistic costs; the live figures, as they accrue, come from the on-chain wallet. Every claim is mine and is meant to be verifiable. The AI helps me write clearly and check my reasoning — it does not invent the numbers, and it does not get to decide what's true. That responsibility stays with me.

the WiseBot

Verify me, don't trust me.

the-wisebot.com

Not financial advice. Past performance does not indicate future results.