REVIEW OF TRADING SYSTEMS · FULL ANALYSIS

One Look at Ichimoku

Ichimoku Kinko Hyo is the most beautiful indicator in retail technical analysis — five lines, a shaded cloud, a system that looks like knowledge. Tested honestly on sixteen years of Bitcoin, its famous constants turn out to be Saturdays, one of its five lines contains literally zero information, its entire measured edge is six trades — and using machine learning to "improve" it is arithmetically guaranteed to lose money.

by **David Kwok Ho Chan**, with AI Sifu Claude · measured on daily BTC, 2014–2019 (2020+ sealed)

Ichimoku means **one look**.

The full name — 一目均衡表, *ichimoku kinko hyo* — translates as something like *the table of equilibrium prices at a glance*. It was built in the 1930s by a Tokyo newspaper writer named Goichi Hosoda, who published under the pen name **Ichimoku Sanjin**: "a glance of a mountain man." The promise is right there in the branding. Five lines, one shaded cloud, and the market's whole condition legible in a single glance.

It is, I think, the most beautiful thing in retail technical analysis. It is also the perfect specimen for this series, because beauty and knowledge are exactly the two things a simulacrum invites you to confuse.

So: one honest look. It has a shape — an intro that says where I stand, **three findings** that each deflate the last, and a conclusion about what's actually left.

> **One — The machine is simpler than it looks.** Strip the mystique and Ichimoku is a moving-average system built on a payroll calendar, with one of its five lines carrying zero information.
>
> **Two — The edge was six trades.** The backtest "works," but the entire result is a handful of avoided bear markets and one catastrophically missed parabola.
>
> **Three — You can't filter a fat tail.** The obvious fix — use machine learning to keep the good trades — is arithmetically guaranteed to lose money here.

The receipts (this applies to me first)

This series has one rule: the test for any method is **does the edge survive out-of-sample, with real costs, and no hindsight?** Which means I have to tell you where I stand relative to my own data before I tell you anything about the cloud.

Everything below is measured on **daily Bitcoin, 2014-01-01 to 2019-12-31** — 2,191 bars. The period from 2020 onward is **sealed**. I have not run a strategy backtest on it, and I'm not going to until I have a hypothesis worth spending it on.

And I owe you a **degrees-of-freedom ledger**, because "I tested it" means nothing without knowing how many times I looked:

> Two parameter sets. Two position variants (long/flat and long/short). Four cost scenarios. One yearly decomposition. Six candidate features. All on the same in-sample slice.

That is a lot of looking. Nothing here is confirmation. It's the *generation* of a hypothesis, and I'll say plainly at the end which single pre-registered test would actually settle it. If that framing feels like excessive throat-clearing, consider that its absence is precisely what makes most published backtests worthless. The honesty isn't a disclaimer bolted onto the analysis. It **is** the analysis.

(None of this is advice. It's a methodology exercise on a public price series.)

One — The Machine Is Simpler Than It Looks

The five lines

Every source garbles these differently, so here they are from the primary text I worked from — Ken Muranaka's write-up in *Stocks & Commodities* V.18:10, one of the few places the original formulas appear in English with the construction spelled out.

LINE	JAPANESE	FORMULA	PLOTTED
Turning line	**Tenkan-sen**	(HH + LL) / 2 over **9**	at today
Standard line	**Kijun-sen**	(HH + LL) / 2 over **26**	at today
1st preceding span	**Senkou Span A**	(Tenkan + Kijun) / 2	shifted **+26 forward**
2nd preceding span	**Senkou Span B**	(HH + LL) / 2 over **52**	shifted **+26 forward**
Delayed line	**Chikou Span**	today's close	shifted **−26 back**

The band between Span A and Span B is the **kumo** — 雲, cloud. Above the cloud, Muranaka writes, "the sun is shining and it would be a time to buy." Below it, "it's raining." Note what's *not* in that table: the **open**. Ichimoku is built entirely from highs, lows, and closes.

Here's the first uncomfortable thing. Muranaka states the system's real nature himself, twice:

> "By and large, the calculations for ichimoku charts are similar to moving average techniques but use the historical highs and lows rather than a series of closing prices."
>
> "An ichimoku chart is a trend-following system with an indicator similar to moving averages."

He then does what a good analyst does and **plots the control** — the same gold contract with a plain 9-day and 26-day moving average — and captions it: *"This shows that an ichimoku chart is a form of trend-following system."*

So the novelty isn't the lines. Muranaka is explicit that it's the *plotting convention*: "the strategy to time-shift the trendlines to the past for the delayed line and to the future for the preceding lines." That's genuinely clever information design. One chart, three questions answered, one glance. But **information design is not edge.** Three specific things go wrong underneath it.

Wrong thing #1: the magic numbers are Saturdays

9, 26, 52. These get treated as though they were handed down — Fibonacci-adjacent, harmonic, discovered. Here is where 26 comes from, in Muranaka's own words:

> "The time span of 26 days would correspond to the number of business days (**Saturdays included**) in one month when this charting method was devised and tested."

Twenty-six is the number of trading days in a month **when the Tokyo exchange traded six days a week.** The most venerated constant in retail technical analysis is a payroll calendar from before the weekend was invented. Muranaka, to his credit, says the obvious thing out loud: "The time spans of nine, 26, and 52 may be changed for the current markets, as securities are not currently traded on Saturday."

This matters more than a trivia point, because Bitcoin trades **seven** days a week, 365 days a year. The faithful translation of Hosoda's *intent* — a week and a half, a month, two months — is not 9/26/52. It's **10/30/60**. I tested both, and the calendar-corrected version was better in every single scenario. That's one comparison, not a proof — but it's the version with an argument behind it, so it's the version I use below.

While we're at it: Hosoda used **midpoints** of highs and lows rather than means, and the received explanation is that he "believed midpoints reflect price characteristics better than averages." Muranaka offers a drier alternative: *"It is likely that taking midpoints were computationally less of a burden than taking averages."* A midpoint of 26 bars is one MAX, one MIN, one division. A 26-bar mean is 26 additions. By hand, daily, across many markets, the midpoint is roughly ten times cheaper. **The signature construction of Ichimoku is very plausibly a compute optimization that got promoted to a theory of market structure.** That move — implementation constraint mistaken for principle — recurs throughout this series.

Wrong thing #2: one of the five lines contains zero information

Look again at Senkou Span A:

$$\text{Span A} = \frac{\text{Tenkan} + \text{Kijun}}{2}$$

It is an algebraic function of two lines you already have. It cannot contain anything they don't. I checked it numerically rather than trusting the algebra, comparing the plotted Span A against a reconstruction from Tenkan and Kijun:

```
max | span_a − shift((tenkan + kijun)/2) |  =  0.0000000000
```

Exactly zero. Which means the celebrated five-line "confluence" setup — the one where all five lines agree and you're supposed to feel very confident — is not five independent witnesses. Strip the redundancy and you have:

- a 10-bar midpoint
- a 30-bar midpoint
- a 60-bar midpoint

- the sign of 30-bar momentum

Three nested midpoints of the same series, plus a momentum bit. **Confluence between correlated indicators isn't confirmation. It's double-counting.** If you take one thing from this piece, take that one — it generalizes far past Ichimoku, and it's why "I use six indicators and wait for them all to line up" is a much weaker statement than it sounds.

Wrong thing #3: the Chikou span is a lookahead landmine

This one is a genuine public service, because I suspect it is silently wrong in a large fraction of the Ichimoku systems on the internet. The Chikou span is today's close, plotted 26 bars **in the past**. The standard implementation is therefore a negative shift:

```
chikou = close.shift(-26)      # correct for PLOTTING
```

Muranaka's own spreadsheet sidebar flags exactly what this does:

> "Column G, the delayed line, will end in a series of zeroes because **it looks ahead in the data to see what the values are in the future.** This formulation, while unusual to Western eyes, is correct."

Correct **for drawing a chart**. Catastrophic if you feed it to a signal. Because if your backtest compares `chikou[t]` to `close[t]`, you have compared the close *twenty-six bars from now* to today's close. You have not built an indicator. You have built a time machine. I measured how much that's worth — directional hit rate against the forward 30-bar return:

VERSION	HIT RATE
legitimate — $sign(\text{close}_t - \text{close}_{t-30})$	**53.0%**
lookahead — $sign(\text{chikou}_t - \text{close}_t)$	**100.0%**

Not "somewhat better." **Perfect.** A flawless oracle, on every single bar. That is what a sign error buys you, and it is why a certain kind of Ichimoku backtest produces an equity curve that looks like a rocket launch. The honest content of the Chikou span is much smaller, and Muranaka says so plainly: "This is simply a comparison of the current prices with the prices as of a month ago, and **the quick comparison is the only use of the delayed line.**" It's a 30-period momentum sign. One bit.

> **Rule: before you believe any Ichimoku result, grep the code for `shift(-`.**

And the cloud itself? A random band would do as well

The kumo's headline claim is that it "defines support or resistance." Adam Grimes ran the control experiment that nobody running an indicator ever runs — he hid the price bars, drew lines on the chart at random, and looked at what happened when price reached them:

> "Almost any random line drawn in the market will function as believable support or resistance."

That's the correct null hypothesis for the cloud, and it's a demanding one. On Bitcoin the kumo's median thickness is **6.2% of price**, with a 90th percentile of 18%. That is not a level. It's a *region*, one that price is mechanically guaranteed to interact with constantly. Screenshots of price "respecting the cloud" are therefore free. Before any of them count as evidence, the cloud has to beat a random band of matched width — and on gold, the market Hosoda designed this for, that band would have been perhaps 1–2% wide. On crypto it's a blunt instrument.

So much for the machine: the lines are a moving average in costume, one of them is empty, one is a lookahead trap, and the cloud is a random band with good branding. **The question that actually matters is whether the thing makes money anyway.**

Two — The Edge Was Six Trades

Enough deflation. Does it *work?* Rules: go long when the daily close is above the cloud, flat otherwise. Signals act on the next bar — no same-bar execution. Costs charged from the first trade. Perpetual futures, because perps never expire, which means no roll gaps contaminating those 60-bar MAX/MIN inputs.

	MULTIPLE	CAGR	SHARPE	MAX DRAWDOWN
Buy & hold	9.34x	45.1%	0.87	**−83.3%**
Ichimoku, perp, neutral funding	**19.45x**	64.1%	**1.23**	**−48.9%**
Ichimoku, perp, *warm* funding	10.63x	48.3%	1.03	−50.7%
Ichimoku, **retail spot** (3.2% round trip)	**2.65x**	17.7%	0.57	−66.8%

Double the money and half the drawdown. On its face, a win. Three things demolish that reading.

The venue matters more than the signal

Look at the last two rows. The **identical rules** return 19.45x on a perp and **2.65x on retail spot** — against 9.34x for doing nothing at all. The same signal is a winner on one venue and *actively destructive* on another, purely because retail spot fees are roughly forty times perp fees. And funding is not a detail. Dropping from neutral to warm funding takes 19.45x down to 10.63x — barely past buy-and-hold. Worse, on a perp **funding is highest exactly when the trend is strongest**: you pay the most precisely when the cloud is most bullish. The cost is positively correlated with the signal. A flat funding assumption is optimistic by construction. Before you ask whether a strategy works, ask **where.**

The Sharpe improvement is inside the error bars

0.87 → 1.23 looks decisive. Over six years with roughly 35 round trips, the standard error on an annualized Sharpe is about

$$SE \approx \sqrt{\frac{1 + SR^2/2}{T}} \approx 0.5$$

So it's 1.23 ± 0.5 against 0.87 ± 0.5. Thoroughly overlapping. **Statistically indistinguishable from luck.** The drawdown improvement is real — but it's *mechanical*, not clever: you're flat half the time

in a violently volatile asset, so of course your drawdown shrinks. You bought that with forgone upside, which makes this a **risk transformation, not an edge.**

And then the year-by-year

This is the part that changed my mind.

YEAR	BUY & HOLD	ICHIMOKU	EXCESS
2014	−58.5%	−3.8%	**+54.7%**
2015	+34.6%	+28.9%	−5.6%
2016	+123.8%	+73.8%	−50.1%
2017	+1356.4%	+421.5%	**−934.9%**
2018	−73.3%	−24.2%	**+49.1%**
2019	+92.3%	+128.3%	+36.0%

Decomposing the total log excess of +0.734: **2018 contributes +1.04, 2014 contributes +0.84, and 2017 contributes −1.03.** The entire result is **two avoided bear markets, paid for by one catastrophically missed parabola.** Effective sample size for the favourable effect: **two.**

Look closer at 2017. The system was long **91% of the time** and still captured only 421% of a 1356% year. The 9% it sat out was concentrated in exactly the wrong places — sharp bull-market dips that knocked price under the cloud right before the vertical recoveries. In a compounding series, the best days *are* the return, and a filter whose input is a 30-bar midpoint will always re-enter late. Now the trade level — 35 trades:

METRIC	VALUE
win rate	**45.7% (16W / 19L)**
average win	**+43.5%**
average loss	**−4.7%**
payout ratio	**9.28**
expectancy per trade	**+17.3%**
median trade	**−0.07%**
Wilson 95% CI on win rate	**[30.5%, 61.8%]**

You lose more often than you win. The typical trade is a coin flip that loses a hair. And your win rate is known to ±15 points, which is the honest precision available on any "probability" you'd quote. Then the number that reframes everything:

> **Six trades out of thirty-five produce +623.9%. The total is +606.3%. The other twenty-nine trades collectively lose 17.6%.**

Remove the top trade and expectancy falls from +17.3% to +11.6%. Remove the top three, +4.5%. Remove the top six: **−0.6%.** So what are you actually betting on? Not Ichimoku. You are betting that

Bitcoin keeps producing 50–200% multi-month runs. Ichimoku is a delivery mechanism for that bet, and a leaky one. The regime assumption *is* the position — and it's a far shakier assumption in a post-ETF, institutionally-hedged market than it was in 2016. Jack Schwager's interviewees describe this shape from the inside:

> "you have to minimize your losses and try to preserve capital for **those few instances when you can make a lot in a very short period of time.** What you can't afford to do is throw away your capital on suboptimal trades."

If the whole edge is six trades, the obvious next move is to get better at spotting them. That's exactly the move that can't work.

Three — You Can't Filter a Fat Tail

The natural next thought — I've never met a quantitatively-minded person who doesn't have it within about ninety seconds — is: *fine, so use machine learning to tell the six good trades from the twenty-nine bad ones.* That idea has a name. Marcos López de Prado calls it **meta-labeling**: let the primary model choose the side, and train a secondary model to decide whether to act. It's a real, respected technique. It cannot work here. And the reason is not "not enough data." It's **arithmetic.**

The economics of a perfect filter

6 tail trades sum to	**+623.9%**
29 non-tail trades sum to	**−17.6%**
So a **flawless** filter dropping all 29 losers gains	**+17.6% / 6yr**
Cost of **one** false negative on the biggest winner	**−213.5%**

The upside of a perfect filter is 12.1× smaller than the cost of missing a single tail trade. Sit with that. The thing you're trying to buy — avoiding the small losers — is nearly worthless, because the small losers nearly cancel out. The thing you're risking is everything. I ran the Monte Carlo across realistic classifier performance, where *recall* is the fraction of the six tail trades the filter correctly keeps:

TAIL RECALL	EXPECTED TOTAL	VS NO FILTER	P(FILTER HELPS)
100%	623.9%	**+17.6%**	100%
95%	594.6%	**−11.7%**	**74.4%**
90%	561.7%	−44.6%	53.4%
80%	501.1%	−105.2%	26.7%
60%	376.5%	−229.8%	5.2%

You need essentially 100% recall on a six-example class just to break even. And now look at the bolded row, because it is the most instructive line in this entire analysis. At 95% recall — which

would be an *outstanding* classifier — the filter is **negative in expectation** while **improving the outcome 74.4% of the time.**

That is a machine for fooling yourself. Three times out of four it looks like it's working. You'd ship it, you'd defend it in a meeting, you'd have a chart proving it, and it would slowly cost you money. Negative skew doesn't announce itself; it waits. So, generalized:

> **For fat-tailed payoffs, precision-improving machine learning destroys value.** It optimizes the thing that doesn't matter — avoiding small losses — at the cost of the only thing that does: never missing the tail.

More data doesn't fix this. It fixes a *different* problem (6 positive examples against 6 candidate features, when the rule of thumb wants 10–20 events *per feature* — a fifteen-fold overfit-budget violation). But the filter economics are **structural.** With infinite data, being more selective is still the wrong objective for this payoff shape.

The consolation prize is real

I expected the durations of these trends to be **memoryless** — geometrically distributed, meaning trend age tells you nothing, meaning nothing was learnable at all. They aren't:

TRADE HAS SURVIVED	HAZARD (CHANCE IT ENDS HERE)
1–3 bars	**0.40**
4–10 bars	**0.33**
11–30 bars	**0.21**

Surviving the first ten bars roughly halves your hazard rate. That's genuine trend persistence, measured rather than assumed. (The longer buckets are finite-window artifacts; my formal test against the geometric distribution was fit on the same data it tested, so treat the hazard table as the honest evidence and not the p-value.) But notice **when** that information arrives. Not at entry — *after*. You cannot use it to pick trades. You can use it to size them. Which flips the whole prescription:

1. **Never filter. Modulate size, with a floor.** A binary go/no-go incurs the −213.5% cliff. A continuous sizing function merely *undersizes* a tail winner. This is what meta-labeling is actually for — it outputs size, not side.
2. **Scale in on the measured hazard drop.** Enter small; add after the trade survives ~10 bars. Don't predict which signal is solid — let the market reveal it, and pay for the information with a smaller initial bet. Since the median trade dies inside a week, this is also the cheapest possible way to hold the losers.

And a warning about the other obvious lever: naive Kelly sizing on these numbers recommends **39.9% of equity per trade.** On a distribution whose median trade is −0.07% and whose per-trade standard deviation is 47.9%, that isn't optimal, it's a coin flip on ruin. Kelly assumes you *know* the payout ratio; here it's estimated from sixteen wins dominated by two outliers, and Kelly is savagely sensitive to overestimating it. Quarter-Kelly to tenth-Kelly — 4% to 10% — is the defensible range.

In one look: what survives, and the test that would settle it

Three findings, one direction: the machine is a costumed moving average, its edge is six trades, and you can't machine-learn your way to those six. After all of it, Ichimoku is not worthless. It's just much smaller than it looks:

- **Cloud position as a regime filter.** Objective, non-repainting, zero fitted parameters. Its value is as a *gate on someone else's entry*, not as an entry — and it isn't really an Ichimoku claim at all. It's the trade-with-the-trend claim, which has independent support.
- **Kumo thickness as a volatility-compression proxy.** When the fast midpoints converge on the 60-bar midpoint, the market has gone nowhere for two months. That's a genuine state variable, and it's a *volatility* read rather than a support read.
- **Kijun-sen as a mechanical trailing reference.** Unambiguous, no discretion. Drawdown control is where the demonstrated value actually lives.
- **The displaced cloud as a commitment device.** The forward cloud predicts nothing — it's a stale midpoint wearing future dates. But because it's fully determined 30 bars ahead, your trigger level is *knowable in advance and cannot be moved after the fact.* Worth something to a discretionary trader, just not what the tutorials claim.

And what goes in the bin: the forward cloud as forecast, the "kumo twist" as a turning-point signal, Chikou "confirmation" as anything beyond one momentum bit, five-line confluence as high-conviction evidence, and 9/26/52 as sacred.

One last piece of context. John Murphy's *Technical Analysis of the Financial Markets* is the standard Western reference — six hundred pages, a full chapter on Japanese candlesticks, and **not one page on Ichimoku.** For a system invented in the 1930s and published in 1968, that's a long silence. Muranaka's 2000 article is explicitly framed as a *revival* — "The Mountain Man Is Back." Ichimoku's popularity curve tracks charting software and the retail-FX boom considerably better than it tracks any accumulation of evidence.

Which is the whole thesis of this series. The cloud is beautiful. It renders gorgeously. It makes a chart look like a system, and a system look like knowledge. That gap — between looking like knowledge and being knowledge — is where the money goes to die. *Ichimoku* means one look. It turns out one look is exactly the problem.

The one test that would settle it

Everything above is in-sample and exploratory. So here's the pre-registered specification, written down before the sealed data is touched, which is the only thing that makes pre-registration mean anything:

```
10/30/60 params, long/flat only, cloud-position gate only,
perpetual futures fees plus realistic (not neutral) funding,
evaluated once on 2020-2026, reported whatever it says.
```

Committed to version control before the holdout is opened. My honest prior: the calendar-corrected parameters and the plain moving-average control will be statistically indistinguishable, and the regime filter will carry whatever value shows up. But the question upstream of all of it is the one I'd actually attack first — **has Bitcoin's fat right tail compressed since 2020?** Because

everything here rests on those six events recurring. If the parabolas are gone, so is the strategy, and no amount of parameter tuning will tell you that.

Review of Trading Systems · One Look at Ichimoku · full analysis. Companion to the highlight at 149.28.225.2/posts/ichimoku-one-look/. Not investment advice — a methodology exercise on a public price series.