

# Review — Nison, *Japanese Candlestick Charting Techniques*

Source: Steve Nison, 1991. 316 pages. Ingested to `retail_trading.db` as

`Nison_Japanese_Candlestick_Charting_Techniques` (313 chunks, clean text layer, no OCR).

Test data: BTC daily, 2014-01-01 → 2019-12-31 (2,191 bars). 2020+ holdout sealed.

---

## 1. Summary

---

Nison's 1991 book introduced Japanese candlestick charting to Western markets, and it succeeded so completely that candlesticks are now the default rendering on essentially every exchange and charting platform on earth. The book has two halves.

**Part 1 — the vocabulary.** Roughly forty named patterns built from four numbers per session:

real body= $|C-O|$  upper shadow= $H-\max(O,C)$  lower shadow= $\min(O,C)-L$

***Note:** White body if  $C \geq O$ , black if  $C < O$ . Every pattern in the book is a boolean predicate over these quantities and their lags. Nison never writes them as formulas; the formalisation below is mine.*

Representative examples, formalised:

- **Hammer / hanging man:**  $(\min(O,C)-L) \geq 2|C-O| \wedge (H-\max(O,C)) \approx 0$
- **Bullish engulfing:**  $C_{t-1} < O_{t-1}, C_t > O_t, C_t \geq O_{t-1}, O_t \leq C_{t-1}$
- **Piercing:**  $O_t \leq L_{t-1} \wedge O_{t-1} + C_{t-2} \leq C_t \leq O_{t-1}$
- **Window (gap):**  $L_t \geq H_{t-1}$  or  $H_t \leq L_{t-1}$

***Note:** A hammer and a hanging man are the same line. Only the surrounding trend distinguishes them. The shape carries no information; the context assigns all of it.*

**Part 2 — the doctrine.** The *Rule of Multiple Technical Techniques*, borrowed from Sklarew: *"the more technical indicators that assemble at the same price area, the greater the chance of an accurate forecast."* Condensed in Nison's DISCIPLINE mnemonic as **"Indicators — the more the better."**

**What this review did.** Formalised eight patterns, located every occurrence in six years of daily Bitcoin, and measured forward 5-bar returns against appropriate controls — including a sweep over seven definitions of "trend" and a regime-matched control.

---

## 2. Strengths

---

**The visual encoding is a genuine contribution.** Candlesticks compress open-close direction into a single pre-attentive cue — filled versus hollow — that bar charts bury. Universal adoption is not an accident. This is real information design.

**Nison is unusually honest about the method's limits.** Most technical authors write with false precision. He does the opposite:

*"As with all charting methods, candlestick chart patterns are subject to the interpretation of the user. This could be viewed as a limitation."*  
*"Remember, the following interpretations are subjective... There are no concrete rules."*

**He debunks his own category.** On reversal patterns:

*"the term 'reversal pattern' is somewhat of a misnomer... This rarely happens. Trend reversals usually occur slowly, in stages, as the underlying psychology shifts gears."*  
**Note:** *This is correct and it contradicts how candlesticks are almost universally taught. Credit where due.*

**The risk-management material is the best part of the book** and is independently correct:

*"A stop should be placed at the time of the original trade; this is when one is most objective. Stay in the position only if the market performs according to expectations."*

**He insists on confirmation.** Repeatedly tells the reader to wait for the next session before acting. Whatever the patterns are worth, this advice reduces turnover and cost.

**The scholarship is careful.** Japanese terminology, etymology, and provenance are preserved rather than flattened — *yorikiri* for belt-hold, the sumo metaphor intact. As a work of translation and transmission the book is excellent, and that was the assignment.

**It is a usable taxonomy.** Even granting everything in §3, having shared names for recurring price shapes has communication value.

---

### 3. Weaknesses

---

#### 3.1 No statistics anywhere

A targeted query across all 313 chunks for success rates, sample sizes, hypothesis tests, or backtests returns Fibonacci retracement tables and the stochastic %K formula. There is no *n*, no *p*, and not one frequency count in 316 pages.

All evidence is **hand-picked chart exhibits** — hundreds of windows where the pattern did

what the caption says. This is selection of confirming cases, a procedure that cannot produce a negative result and therefore carries no evidential weight. Grimes' random support/resistance experiment is the relevant control: convincing-looking exhibits can be generated from lines drawn at random.

#### 3.2 The patterns add nothing measurable

**Regime baselines (forward 5-bar):** all bars **+0.870%**, downtrend bars **+0.089%**, uptrend bars **+1.565%**.

**Note:** *The trend regime alone explains a 1.48pp gap. Any pattern that fires in only one regime inherits that gap for free. Comparing patterns to all bars therefore*

*confounds the candle with the regime — and that confound is the entire apparent effect in most published candlestick studies.*

**Regime-matched control** — each pattern versus other bars in the *same* regime:

| pattern           | n  | vs ALL bars (naive)                 | vs SAME regime | 95% CI         |
|-------------------|----|-------------------------------------|----------------|----------------|
| bullish engulfing | 49 | -0.45%                              | +0.35%         | [-1.80, +2.59] |
| bearish engulfing | 36 | -0.28%                              | -1.01%         | [-3.90, +2.04] |
| hammer            | 37 | -2.96% ( <i>looks significant</i> ) | <b>-2.26%</b>  | [-5.36, +0.68] |
| hanging man       | 25 | -1.22%                              | -1.96%         | [-6.90, +2.84] |
| shooting star     | 29 | +0.90%                              | +0.21%         | [-2.61, +2.99] |

**Note: Every confidence interval spans zero.** The hammer's apparent wrong-way signal in the naive comparison is largely the trend filter, not the candle. Conclusion:

**the candlestick shape adds nothing detectable once the trend is known.** What little predictive content exists lives entirely in the trend context — which Nison requires on every page and defines on none.

### 3.3 The result is robust to the free parameter

Because Nison never defines "downtrend," any test must invent one. Swept across seven definitions (SMA 5/10/20/50, 5-bar and 20-bar returns, and no filter), 43 testable cells:

- **32 of 43 (74%)** underperform the unconditional baseline
- 6 flagged significant versus 2.1 expected by chance — but these are the *naive* comparisons, and §3.2 shows the regime control absorbs them

- **The hammer is negative under all seven definitions**, growing more negative with longer windows (−1.38% at SMA5 → −6.35% at 20-bar return)

***Note:** The direction of the finding does not depend on my arbitrary choice. That parameter is nonetheless an unlogged degree of freedom in every candlestick study ever published, this one included — logged here.*

### 3.4 Several patterns barely occur

Piercing appeared **4 times** and dark cloud cover **5 times** in six years — untestable at almost every trend definition. Patterns given pages of exposition are near-absent from the data.

### 3.5 Unfalsifiability, converted into a virtue

Having admitted the subjectivity quoted in §2, Nison writes:

***"In this sense, subjectivity may not be a liability."***

***Note:** A method with no concrete rules cannot be wrong: a failed hammer becomes a misjudged trend, or not a true hammer, or a missing confirmation. Every outcome is absorbed. Unfalsifiability buys immunity from criticism at the price of any evidence in favour.*

### 3.6 "More indicators the better" is backwards

Candlesticks, RSI, stochastics, trendlines and retracements are all deterministic functions of one OHLC series. Stacking them is asking one witness five times and recording five answers. This is the identical error measured in the Ichimoku review, where Senkou Span A

proved to be *algebraically*  $(\text{Tenkan} + \text{Kijun})/2$  — residual 0.0000000000, exactly zero independent information. **Confluence among correlated indicators is double-counting, not confirmation** — and each added indicator multiplies the search space, so the doctrine is simultaneously a recipe for overfitting.

### 3.7 The method is venue-dependent, unacknowledged

Candlestick semantics rest on the **open**. Ichimoku never uses it (Muranaka: "Open is not used"). In a 24/7 market the open is merely the last tick before midnight UTC:

| quantity                           | value         |
|------------------------------------|---------------|
| median $ O_t - C_{t-1}  / C_{t-1}$ | <b>0.048%</b> |
| median $ C_t - O_t  / O_t$         | 1.458%        |
| ratio                              | <b>0.033</b>  |

***Note:** The open carries ~3% of the information the patterns assume. Consequently the real body  $\approx$  the daily return, and the open-versus-prior-close vocabulary measures a rounding error.*

**Windows: 13 of 2,190 bars (0.59%).** A market with no session break cannot gap. Nison's entire window chapter — plus tasuki gaps, gapping plays, side-by-side white lines — has no referent here.

### 3.8 No transaction costs

Cost appears nowhere as a filter on whether a signal is worth acting on.

## 4. Where Next

**The strongest next test puts my own claim at risk.** §3.7 argues candlesticks fail in crypto *because* the open is degenerate. That is a falsifiable structural hypothesis, and the way to test it is to run the identical harness on an asset with a **real session open** — SPY, or a liquid futures contract with a genuine overnight break.

- If patterns show an edge there but not in crypto → the venue hypothesis is supported.

- If they fail there too → my crypto explanation is superfluous; the patterns simply don't work, and §3.7 is a nice story that explains nothing.

***Note:** This is the cheapest available experiment that can prove me wrong, which is why it goes first.*

### Other directions, ranked:

1. **Cross-asset pooling.** Sample sizes here are 25–96 per pattern, and 4–5 for the rarer ones. Pooling across many liquid perps (allowing for 0.7–0.9 correlation) would put pattern tests in a defensible range.
2. **Test the Rule of Multiple Technical Techniques directly.** Nison's central doctrine is an empirical claim: does stacking correlated indicators improve accuracy, or degrade it? Measure agreement-count versus forward return. Two books now assert it; nobody has tested it.
3. **Intraday timeframes.** If the open is the problem, 4h and 1h candles have no open at all in any meaningful sense — the degeneracy should be *worse*, and measurably so.
4. **The outstanding Ichimoku holdout.** The pre-registered 10/30/60 specification remains unrun, along with the upstream question of whether BTC's fat right tail has compressed post-2020.

**Which system to review next.** Murphy's core Western apparatus — trendlines, support and resistance, head-and-shoulders — is the natural third subject, because both reviews so far have landed on the same verdict: **the regime/trend context does all the work, and the ornate vocabulary layered on top contributes nothing measurable.** Murphy's chart patterns are the biggest remaining body of claims that has never faced that control.

## Relevance map

---

| Project                                             | Connection                                                                                                                                                  |
|-----------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Review of Trading Systems</b> (post 1, Ichimoku) | Same verdict from a second author: only the regime filter survives. Cross-book confirmation of the double-counting finding.                                 |
| <b>fin-mod / X2 search</b>                          | "More indicators the better" is the multiple-comparisons engine already flagged in <code>spike_prediction_method.md</code> — here stated as doctrine.       |
| <b>Fortuna</b>                                      | Candlestick patterns are not worth adding as features: zero measured contribution after regime control, and they would blow the events-per-variable budget. |
| <b>Lab 4 / HMM</b>                                  | The regime confound in §3.2 is exactly what $\gamma_k(i)$ handles properly — soft, fitted regime membership instead of an undefined "down-trend".           |

---

*Not investment advice. Methodology exercise on a public price series. In-sample only; 2020+ holdout unopened.*